У. Прэтт

Цифровая обработка изображений

2




ИЗДАТЕЛЬСТВО
МИР

DIGITAL IMAGE PROCESSING

William K. Pratt

Professor of Electrical Engineering
Image Processing Institute
University of Southern California
Los Angeles, California

A WILEY–INTERSCIENCE PUBLICATION
JOHN WILEY AND SONS
NEW YORK/CHICHESTER/BRISBANE/TORONTO
1978

У. Прэтт

ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

в 2-х книгах

2

Перевод с английского
под редакцией
канд. техн. наук Д. С. ЛЕБЕДЕВА

ИЗДАТЕЛЬСТВО «МИР»
МОСКВА 1982

ББК 32.94
П85
УДК 621.372.542

Прэтт У.

П85 Цифровая обработка изображений: Пер. с англ. — М.: Мир, 1982. — Кн. 2 — 480 с., ил.

Монография крупного американского специалиста в области обработки изображений, наиболее полная в данной области знания.

В русском переводе монография выпускается в двух книгах. В книге 2 описываются алгоритмы улучшения, реставрации и анализа изображений, а также способы цифрового кодирования. При рассмотрении процессов реставрации и кодирования широко используются двумерные линейные преобразования, унитарные преобразования и метод сингулярного разложения матриц.

Предназначена для специалистов, имеющих дело с обработкой изображений, а также для студентов, изучающих методы цифровой обработки изображений.

П $\frac{30401\text{-}186}{041\ (01)\text{-}82}$ 148-81, ч. 1 2405000000

ББК 32.94
П85

Редакция литературы по новой технике

Часть 4

Реставрация и улучшение изображений

В последние годы значительно возрос интерес к электронным, цифровым и оптическим методам обработки изображений с целью повышения их качества. Широкое освещение получили работы, связанные с космическими и биомедицинскими исследованиями. Из числа других применений следует упомянуть аэрофотосъемку и промышленную радиографию.

Повышение качества изображений достигается двумя видами обработки изображений: реставрацией (исправлением) изображений и их улучшением. Под реставрацией обычно понимают процедуру восстановления или оценивания элементов изображения, целью которой является коррекция искажений и наилучшая аппроксимация идеального неискаженного изображения. Для улучшения изображений используется комплекс операций, призванных улучшить восприятие изображения наблюдателем или же преобразовать его в другое изображение, более удобное для машинной обработки.

В гл. 12 описан ряд методов улучшения одноцветных и цветных изображений. В гл. 13—16 рассмотрены модели, используемые при реставрации изображений, описаны пространственные методы реставрации, а также методы поэлементной, спектральной и цветовой реставрации.

Глава 12

УЛУЧШЕНИЕ ИЗОБРАЖЕНИЙ

Процедура улучшения изображений сводится к выполнению комплекса операций с целью либо улучшения визуального восприятия изображения, либо преобразования его в форму, более удобную для визуального или машинного анализа. В системах улучшения изображений не делается попытки приблизить воспроизводимое изображение к некоторому идеализированному оригиналу (такая задача решается при реставрации изображений). Известны случаи, когда искаженное изображение субъективно воспринимается лучше, чем неискаженный оригинал. Примером может служить изображение с подчеркнутыми границами (контурами).

При машинной обработке улучшение изображения тесно связано с задачей извлечения информации. Пусть, например, система улучшения изображений производит подчеркивание границ исследуемого изображения путем высокочастотной фильтрации. Обработанное изображение затем вводится в ЭВМ, которая выполняет операцию прослеживания контура объекта, определяет его форму и размеры. В этом примере система улучшения изображений используется для того, чтобы подчеркнуть важнейшие признаки исходного изображения и, следовательно, облегчить задачу извлечения информации.

В настоящее время мы не располагаем общей теорией улучшения изображений. Это объясняется тем, что до сих пор еще не выработан общепринятый стандарт качества изображения, который мог бы послужить критерием при проектировании систем улучшения изображений. Здесь описан ряд различных методов, с успехом использованных для улучшения условий наблюдения изображений и облегчения машинного анализа.

12.1. ИЗМЕНЕНИЕ КОНТРАСТА

Слабый контраст — наиболее распространенный дефект фотографических и телевизионных изображений, обусловленный ограниченностью диапазона воспроизводимых яркостей, нередко сочетающейся с нелинейностью характеристики передачи уровней. Во многих случаях контраст можно повысить, изменяя яркость

каждого элемента изображения [1, 2]. Рис. 12.1.1, *а* иллюстрирует характеристику передачи уровней, которая требуется для повышения контраста типичных непрерывных малоконтрастных изображений. Коррекцию таких изображений можно осуществить, пользуясь фотографическими методами, однако реализовать произвольную характеристику передачи уровней с высокой точностью обычно затруднительно. В случае цифровых изображений получить требуемую характеристику передачи уровней относительно просто. Однако при отыскании нелинейного оператора следует учитывать ошибки квантования. Пусть исходное изображение, проквантованное на J уровней, имеет ограниченный диапазон яркостей (рис. 12.1.1,*б*). Предполагается, что выходное изображение также квантуется на J уровней, а преобразование линейно. Как видно из рис. 12.1.1, *б*, каждый выбранный выходной уровень является ближайшим к уровню, соответствующему входному уровню. Очевидно, что в диапазоне яркостей выходного изображения некоторые уровни не будут использоваться, поэтому отдельные перепады яркости выходного изображения будут превышать соответствующие перепады исходного изображения. Вследствие этого могут возникнуть заметные ложные контуры. Если выходное изображение квантовать с большим числом уровней, чем входное, то можно получить равномерное размещение выходных уровней и благодаря этому уменьшить эффект появления ложных контуров.

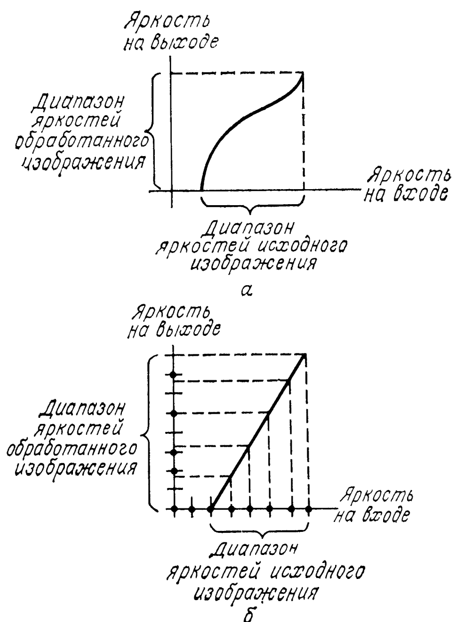


Рис. 12.1.1. Повышение контраста.
а — непрерывное изображение; *б* — цифровое изображение.

Диапазон яркостей выходного изображения, подвергнутого цифровой обработке, может отличаться от диапазона яркостей исходного изображения. Более того, в диапазоне числовых значений яркости обработанного изображения могут оказаться отрицательные значения, которые не могут иметь прямого соответствия с физическими яркостями. Рис. 12.1.2 иллюстрирует два

возможных способа приведения диапазона яркостей выходного изображения в соответствие с диапазоном яркостей исходного изображения. Согласно первому из них, обработанное изображение линейно отображается таким образом, чтобы полностью охватить отведенный ему диапазон яркостей. Второй способ предусматривает ограничение экстремальных значений яркости обработанного изображения максимальным и минимальным порого-

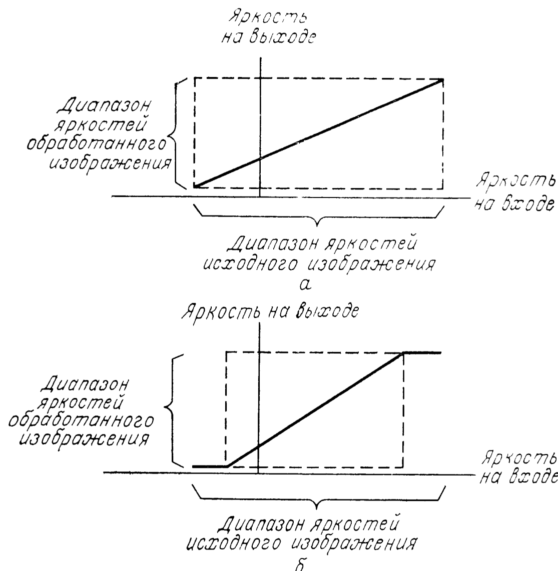


Рис. 12.1.2. Примеры контрастного масштабирования.
а — линейное масштабирование; *б* — линейное масштабирование с ограничением.

выми уровнями. Этот способ часто обеспечивает более высокое субъективное качество изображения, особенно если обработанное изображение содержит относительно мало элементов с превышением уровней ограничения. В программах и системах повышения контраста нередко предусматривается возможность введения ограничения для фиксированного процента значений яркости на обоих краях шкалы яркостей.

Рис. 12.1.3 иллюстрирует результаты повышения контраста одного из изображений со спутника ERTS; здесь же приведены гистограммы распределения яркостей исходного и улучшенных изображений. Рис. 12.1.3, *в* соответствует случаю, когда в качестве уровней ограничения были выбраны экстремальные значения яркости исходного изображения. На рис. 12.1.3, *д* представлено обработанное изображение при выборе уровней ограничения, от-

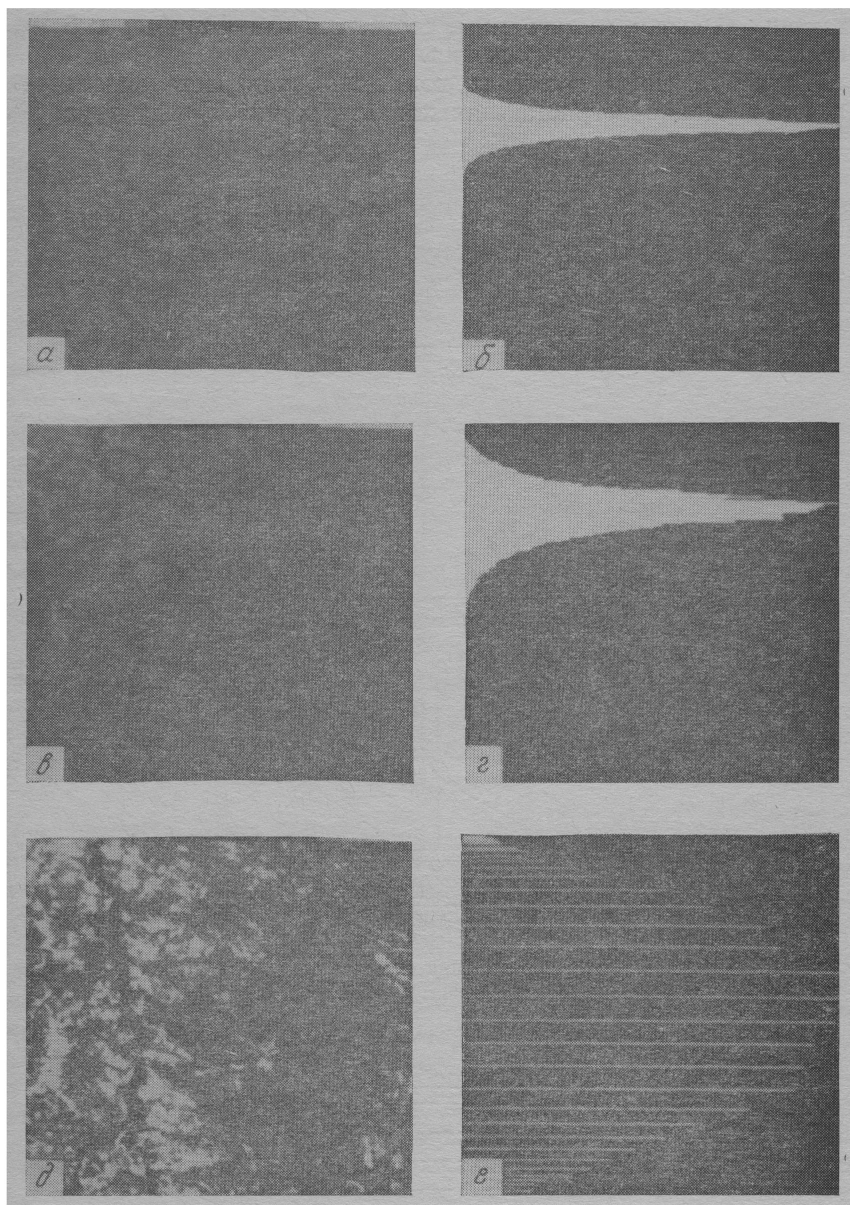


Рис. 12.1.3. Примеры повышения контраста.

а — исходное изображение; *б* — гистограмма распределения яркостей для исходного изображения *а*; *в* — обработанное изображение при уровнях ограничения 43 (мин) и 163 (макс), *г* — гистограмма распределения яркостей для обработанного изображения *в*; *д* — обработанное изображение при уровнях ограничения 62 (мин) и 88 (макс); *е* — гистограмма распределения яркостей для изображения *д*.

секающих «хвосты» гистограммы, соответствующие примерно 5 % элементов исходного изображения.

В рассмотренных примерах повышения контраста использовалась монотонно возрастающая характеристика передачи уров-

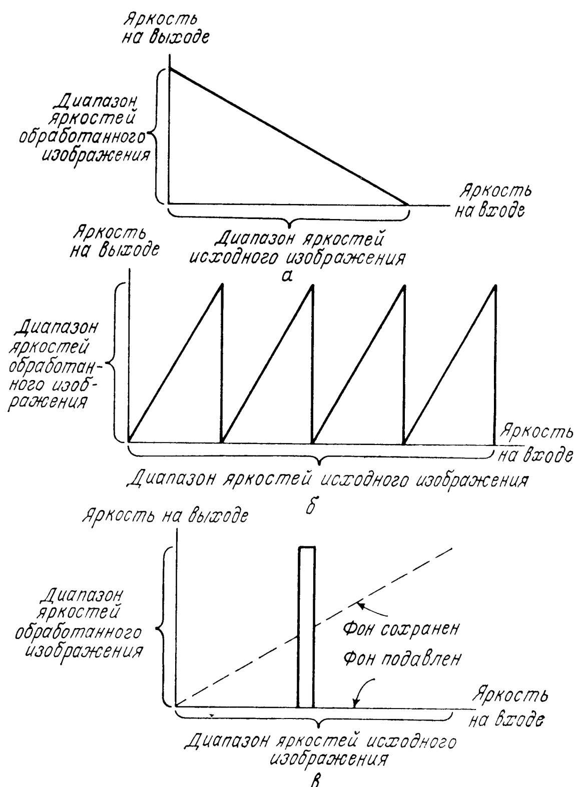


Рис. 12.1.4. Специальные преобразования контраста.

а — обращенное контрастное масштабирование; *б* — пилообразное контрастное масштабирование; *в* — яркостный срез.

ней. В ряде случаев обработки изображений целесообразнее пользоваться монотонно спадающей или немонотонной характеристикой. На рис. 12.1.4, *а* показана характеристика передачи уровней, обеспечивающая обращение шкалы яркостей. Такую характеристику целесообразно применять, например, если дисплей обладает существенной нелинейностью в области черного. Тогда темные участки входного изображения будут переведены в светлые, которые соответствуют линейному участку шкалы яркостей дисплея. На рис. 12.1.4, *б* приведена характеристика передачи уров-

ней, предназначенная для пилообразного контрастного масштабирования. Такое преобразование часто используют с целью получения изображения с широким динамическим диапазоном на экране дисплея, имеющего ограниченный динамический диапазон. Обра-



Рис. 12.1.5. Образцы изображений после специальных преобразований контраста. *а* — исходное изображение; *б* — обращенное контрастное масштабирование; *в* — четырехкратное пилообразное контрастное масштабирование; *г* — белые участки соответствуют яркостному срезу для $16 \leqslant F(j, k) \leqslant 40$.

зец изображения после пилообразного контрастного масштабирования показан на рис. 12.1.5, *в*. Характеристика передачи уровней для яркостного среза изображения приведена на рис. 12.1.4, *в*. Это преобразование позволяет выделить определенный интервал диапазона яркостей входного изображения. Одноцветное изображение после выполнения яркостного среза показано на рис. 12.1.5, *г*. При наличии цветного дисплея появляется возможность отображения в каком-то ярком цвете (например, в красном)

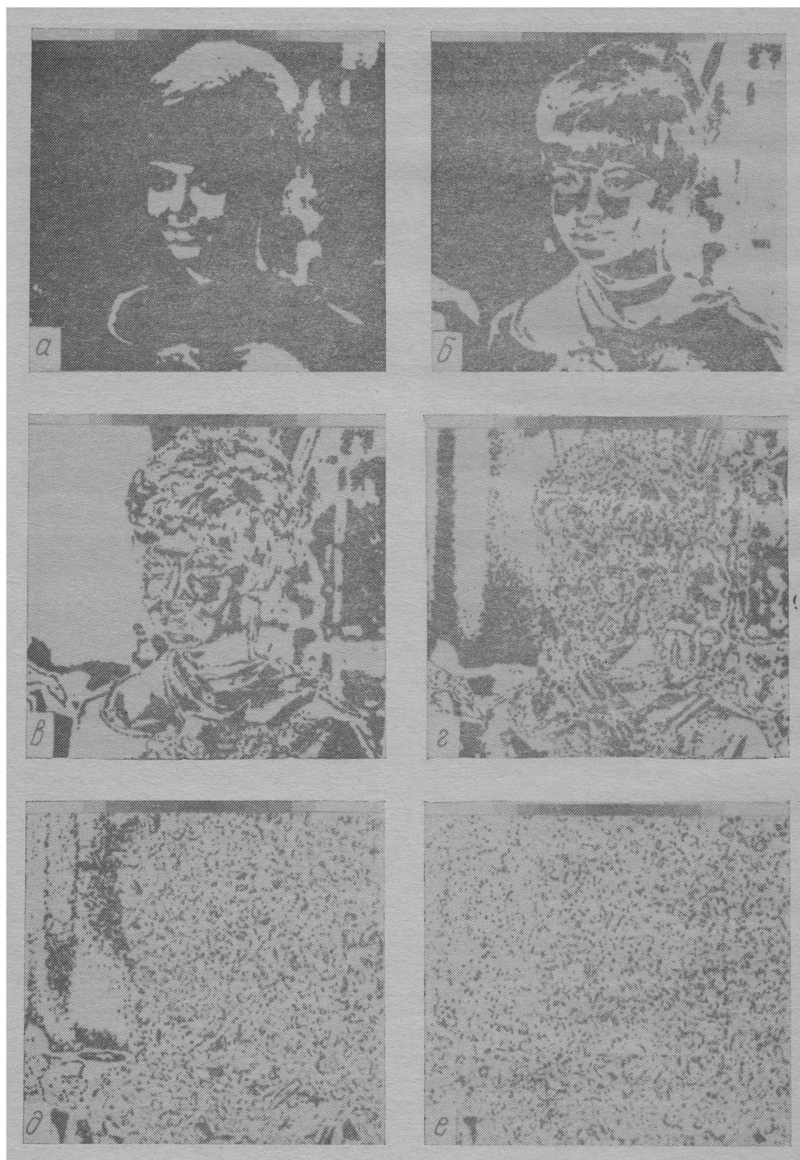


Рис. 12.1.6. Образцы изображений после выполнения поразрядного яркостного среза.
а — разряд 1; *б* — разряд 2; *в* — разряд 3; *г* — разряд 4; *д* — разряд 5; *е* — разряд 6.

участков одноцветного изображения, яркость которых находится в произвольном узком интервале; остальные участки этого изображения отображаются обычным способом. Рис. 12.1.6 иллюстрирует операцию поразрядного среза. Каждое изображение имеет два значения яркости, которые соответствуют данному разряду кодового слова, представляющего яркость. Метод яркостных срезов оказывается исключительно эффективным при визуальном анализе изображений в том случае, когда оператор пользуется дисплеем в диалоговом режиме.

12.2. ВИДОИЗМЕНЕНИЕ ГИСТОГРАММ

Гистограмма распределения яркостей типичного изображения естественного происхождения, подвергнутого линейному квантованию, обычно имеет ярко выраженный перекос в сторону малых уровней; яркость большинства элементов изображения ниже средней. На темных участках подобных изображений детали часто оказываются неразличимыми. Одним из методов улучшения таких изображений является видоизменение гистограммы. Этот метод предусматривает преобразование яркостей исходного изображения, с тем чтобы гистограмма распределения яркостей обработанного изображения приняла желаемую форму. Эндрюс, Холл и другие исследователи [3—5] получили ряд улучшенных изображений путем выравнивания гистограмм распределения, т. е. в каждом случае они добивались равномерности распределения яркости обработанного изображения. Фрей [6] исследовал метод видоизменения гистограмм, который обеспечивал экспоненциальную или гиперболическую форму распределения яркостей улучшенного изображения. Кетчам [7] улучшил этот метод, применив «локальные», подсчитанные по части изображения, гистограммы.

Рис. 12.2.1 иллюстрирует процедуру выравнивания гистограммы в том случае, когда число уровней на выходе составляет половину числа уровней на входе. Здесь $H_F(j)$ — доля элементов входного изображения, квантованная яркость которых соответствует j-му уровню ($j = 1, 2, \ldots, J$). Целью процедуры выравнивания гистограммы является формирование выходного изображения G, нормированная гистограмма которого описывается соотношением $H_G(k) = 1/K$, где $k = 1, 2, \ldots, K$. Алгоритм преобразования реализуется следующим образом. Начиная с наименьшего уровня яркости исходного изображения, объединяют элементы соседних интервалов квантования таким образом, чтобы суммарный результат наименее отличался от величины $1/K$. Все объединенные элементы приводятся к первому новому уровню, который располагается точно в середине первого интервала квантования обработанного изображения. Эту процедуру повторяют для остальных более высоких уровней яркости. При большом

числе уровней квантования исходного изображения, превышающем число уровней квантования улучшенного изображения, удается получить почти равномерную гистограмму. Рассмотренная процедура приводит к некоторому увеличению ошибок квантования. Существует метод выравнивания гистограмм распределения яркостей, который даже при одинаковом числе уровней квантования исходного и улучшенного изображений обеспечивает

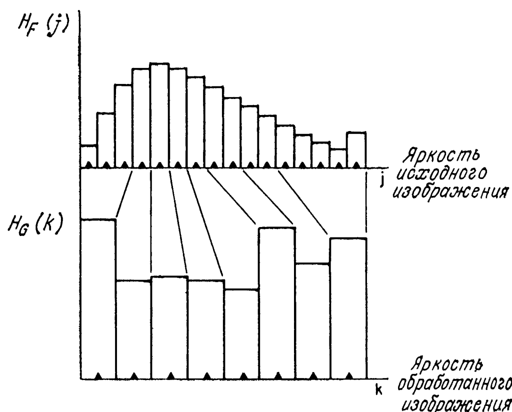


Рис. 12.2.1 Пример приближенного выравнивания гистограммы распределения яркостей, когда обработанное и исходное изображения имеют неодинаковое число уровней квантования.

получение равномерной гистограммы. Этот метод основан на случайном перераспределении элементов изображения при формировании выходных интервалов квантования из входных.

Процедуру видоизменения гистограммы можно рассматривать как монотонное поэлементное преобразование $g_k = T\{f_j\}$ входной интенсивности $f_0 \leqslant f_j \leqslant f_J$ в выходную интенсивность $g_0 \leqslant g_k \leqslant g_K$, в результате которого исходное распределение вероятностей $\Pr\{f_j\}$ переходит в распределение вероятностей $\Pr\{g_k\}$, имеющее желательную форму. Очевидно, что сумма вероятностей всех уровней должна равняться единице. Таким образом,

$$\sum_{j=0}^{J} \Pr\{f_j\} = 1, \qquad (12.2.1а)$$

$$\sum_{k=0}^{K} \Pr\{g_k\} = 1. \qquad (12.2.1б)$$

Кроме того, для любого j должны быть одинаковы исходная и преобразованная функции распределения вероятностей. Другими

словами, вероятность того, что яркость элементов исходного изображения меньше или равна f_j, должна равняться вероятности того, что яркость элементов обработанного изображения будет меньше или равна g_k, где $g_k = T\{f_j\}$:

$$\sum_{n=0}^{k} \Pr\{g_n\} = \sum_{m=0}^{j} \Pr\{f_m\}. \qquad (12.2.2)$$

Сумма в правой части равенства определяется распределением вероятностей для исходного изображения. В случае конкретного изображения это распределение заменяется распределением частот, поэтому можно записать

$$\sum_{n=0}^{k} \Pr\{g_n\} = \sum_{m=0}^{j} H_F(m). \qquad (12.2.3)$$

Обращение этого равенства позволяет получить решение для g_k в зависимости от f_j. В общем случае такое обращение очень трудно или даже невозможно выполнить аналитически, однако численные методы позволяют легко решить эту задачу. Получаемое решение имеет вид таблицы, в которой для каждого входного уровня указывается соответствующий выходной уровень.

Приближенную форму преобразования гистограммы можно получить, если в формуле (12.2.2) дискретные распределения вероятностей заменить плотностями непрерывных распределений вероятностей. Результирующее приближение имеет вид

$$\int_{g_{\text{мин}}}^{g} p_g(g)\,dg = \int_{f_{\text{мин}}}^{f} p_f(f)\,df, \qquad (12.2.4)$$

где $p_f(f)$ и $p_g(g)$ — плотности вероятности f и g соответственно. Интеграл в правой части равенства представляет собой функцию распределения вероятностей $P_f(f)$ входной переменной f. Следовательно, можно записать

$$\int_{g_{\text{мин}}}^{g} p_g(g)\,dg = P_f(f). \qquad (12.2.5)$$

Таблица 12.2.1

Характеристики передачи уровней, используемые при видоизменении гистограмм

Плотность вероятности на выходе	Характеристика передачи уровней
Равномерная $p_g(g) = 1/(g_{\text{макс}} - g_{\text{мин}}), \quad g_{\text{мин}} \leqslant g \leqslant g_{\text{макс}}$	$g = [g_{\text{макс}} - g_{\text{мин}}]\, P_f(f) + g_{\text{мин}}$
Экспоненциальная $p_g(g) = \alpha \exp\{-\alpha(g - g_{\text{мин}})\}, \quad g \geqslant g_{\text{мин}}$	$g = g_{\text{мин}} - (1/\alpha) \ln[1 - P_f(f)]$
Рэлея $p_g(g) = \frac{g - g_{\text{мин}}}{\alpha^2} \exp\left\{-\frac{(g - g_{\text{мин}})^2}{2\alpha^2}\right\}, \quad g \geqslant g_{\text{мин}}$	$g = g_{\text{мин}} + \left[2\alpha^2 \ln \frac{1}{1 - P_f(f)}\right]^{1/2}$
Степени 2/3 $p_g(g) = (1/3)\left[g^{-2/3} / \left(g_{\text{макс}}^{1/3} - g_{\text{мин}}^{1/3}\right)\right]$	$g = \left(\left[g_{\text{макс}}^{1/3} - g_{\text{мин}}^{1/3}\right]\left[P_f(f)\right] + g_{\text{мин}}^{1/3}\right)^3$
Гиперболическая $p_g(g) = 1/g\,[\ln(g_{\text{макс}}) - \ln(g_{\text{мин}})]$	$g = g_{\text{мин}}\,[g_{\text{макс}}/g_{\text{мин}}]^{P_f(f)}$

Примечание. Функция распределения вероятностей $P_f(f)$ исходного изображения аппроксимируется гистограммой распределения $P_f(f) \approx \sum_{m=0}^{j} H_F(m)$.

В частном случае, когда требуется, чтобы преобразованная плотность вероятности была равномерной, т. е.

$$p_g(g) = 1/(g_{\text{макс}} - g_{\text{мин}}) \tag{12.2.6}$$

в пределах $g_{\text{мин}} \leqslant g \leqslant g_{\text{макс}}$, характеристика передачи уровней, обеспечивающая выравнивание гистограммы, имеет вид

$$g = [g_{\text{макс}} - g_{\text{мин}}]\, P_f(f) + g_{\text{мин}}. \tag{12.2.7}$$

В табл. 12.2.1 представлены некоторые разновидности гистограмм обработанных изображений и соответствующие им характеристики передачи уровней.

Фрей [6] предложил метод гиперболизации гистограмм (см. табл. 12.2.1), который поясняет рис. 12.2.2. Согласно этому методу, гистограмма исходного изображения видоизменяется так, что плотность вероятности выходного изображения принимает

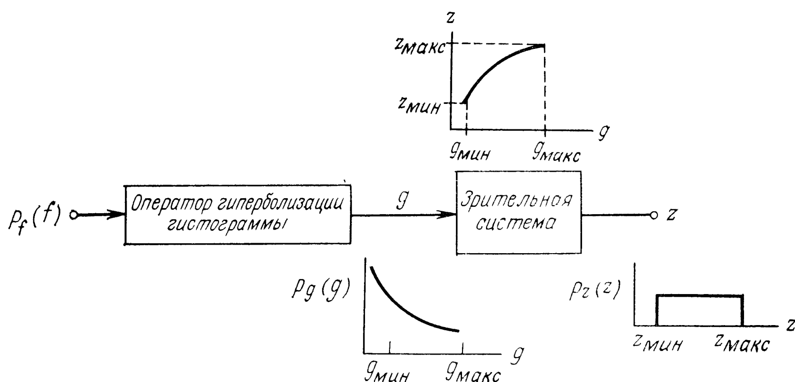


Рис. 12.2.2. Гиперболическое видоизменение гистограммы.

гиперболическую форму. Если теперь допустить, что в зрительной системе выходной сигнал фоторецепторов пропорционален логарифму или кубическому корню входной яркости, то результирующая плотность вероятности окажется равномерной. В данном случае выравнивание гистограммы производится с учетом свойств палочек и колбочек сетчатки.

На рис. 12.2.3 приведены примеры изображений, улучшенных методом видоизменения гистограмм. Изучение представленных образцов показывает, что степень улучшения субъективного восприятия зависит от характера изображения.

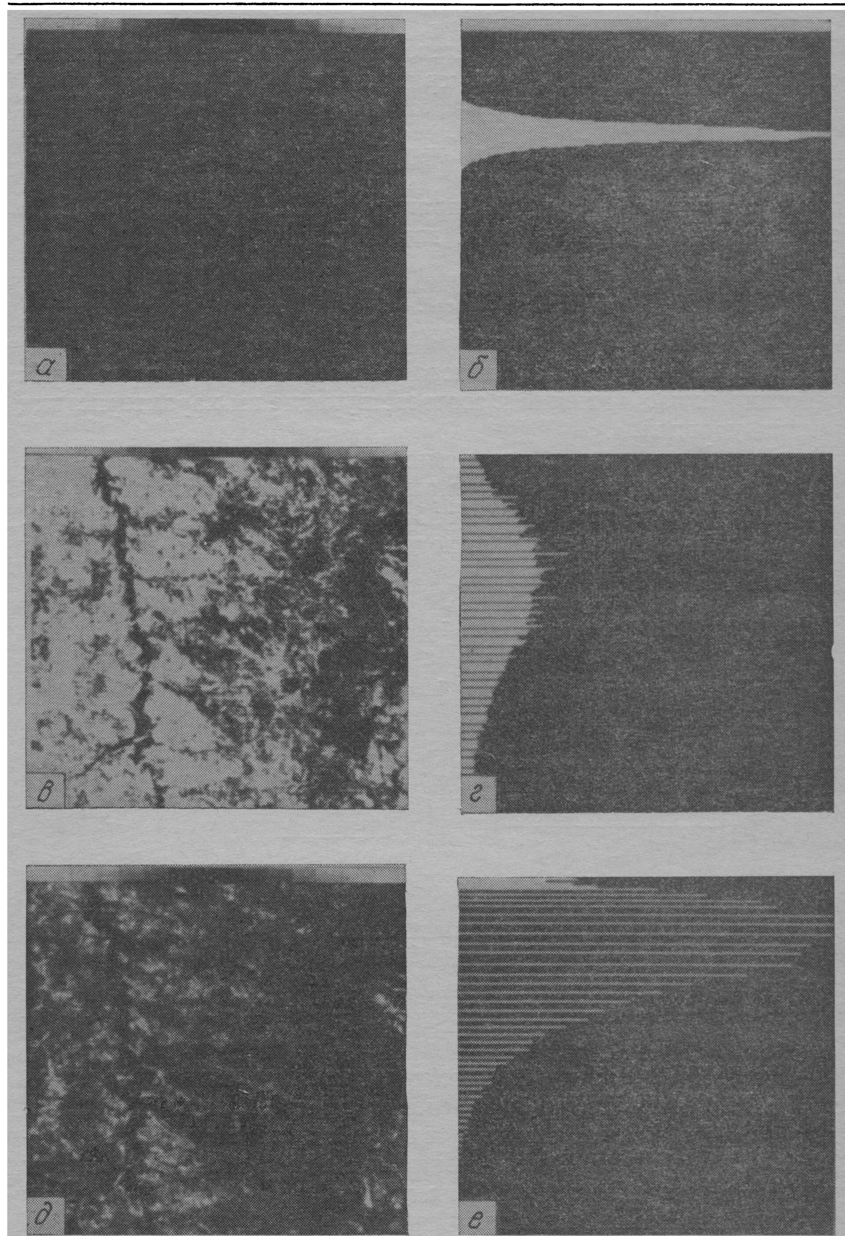


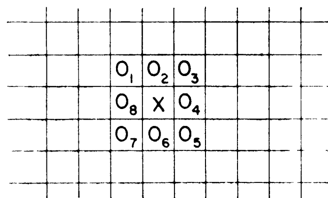
Рис. 12.2.3. Примеры видоизменения гистограмм.

а — исходное изображение; *б* — гистограмма исходного изображения; *в* — изображение после выравнивания гистограммы; *г* — гистограмма улучшенного изображения *в*; *д* — изображение после гиперболического видоизменения гистограммы; *е* — гистограмма улучшенного изображения *д*.

12.3. ПОДАВЛЕНИЕ ШУМОВ

Изображение может повреждаться шумами и помехами различного происхождения, например шумом видеодатчика, шумом зернистости фотоматериалов и ошибками в канале передачи. Их влияние можно минимизировать, пользуясь классическими методами статистической фильтрации (см. гл. 14). Другой возможный подход основан на использовании эвристических методов пространственной обработки.

Шумы видеодатчиков или ошибки в канале передачи обычно проявляются на изображении как разрозненные изменения изолированных элементов, не обладающие пространственной корреляцией. Искаженные элементы часто весьма заметно отличаются от соседних элементов. Это наблюдение послужило основой для многих алгоритмов, обеспечивающих подавление шума [8—11]. Рис. 12.3.1 поясняет простой пороговый метод подавления шума, при использовании которого последовательно измеряют яркость всех элементов изображения. Если яркость данного элемента превышает среднюю яркость группы ближайших элементов на некоторую пороговую величину, яркость элемента заменяется на среднюю яркость. Рис. 12.3.2 иллюстрирует эффективность этого алгоритма применительно к изображениям, переданным посредством ИКМ по двоичному симметричному каналу с вероятностью ошибок 10^{-2}.



Если $\left[X - \frac{1}{8}\sum_{i=1}^{8} O_i\right] > \epsilon$, *то*

$$X = \frac{1}{8}\sum_{i=1}^{8} O_i$$

Рис. 12.3.1. Пример алгоритма подавления шума.

Поскольку шум пространственно декоррелирован, в его спектре, как правило, содержатся более высокие пространственные частоты, чем в спектре обычного изображения. Следовательно, простая низкочастотная пространственная фильтрация может служить эффективным средством сглаживания шумов. В соответствии с методами, рассмотренными в части 3, массив Q размера $M \times M$ выходного изображения формируется путем дискретной свертки массива F размера $N \times N$ исходного изображения со сглаживающим массивом H размера $L \times L$ согласно формуле

$$Q(m_1, m_2) = \sum_{n_1} \sum_{n_2} F(n_1, n_2) H(m_1 - n_1 + 1, m_2 - n_2 + 1). \quad (12.3.1)$$

Сглаживание шума обеспечивается низкочастотной фильтрацией с помощью массива H с положительными элементами. Ниже

приведены сглаживающие массивы трех разновидностей, часто называемые шумоподавляющими масками:

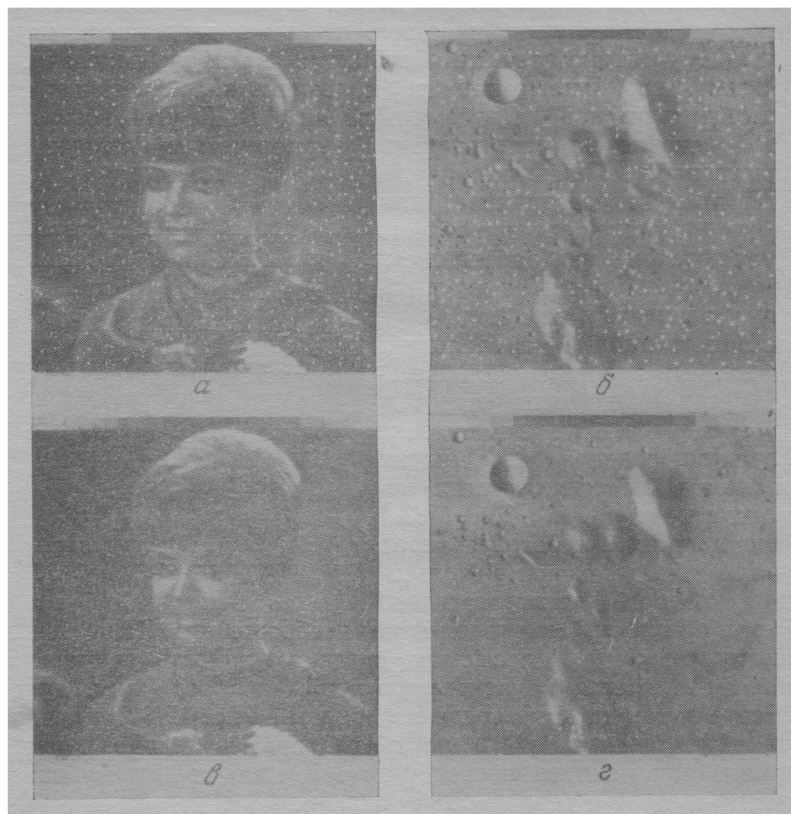


Рис. 12 3.2. Образцы изображений, обработанных методом порогового подавления шума при $\varepsilon = 49$.

а — зашумленное изображение «Портрет»; *б* — зашумленное изображение лунной поверхности, *в* — обработанное изображение *а*, *г* — обработанное изображение *б*

маска 1

$$\mathbf{H} = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \qquad (12.3.2а)$$

маска 2

$$\mathbf{H} = \frac{1}{10}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \qquad (12.3.2б)$$

маска 3

$$\mathbf{H} = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \qquad (12.3.2в)$$

Эти массивы нормированы для получения единичного коэффициента передачи, чтобы процедура подавления шума не вызывала смещения средней яркости обработанного изображения. Эффективность приведенных масок иллюстрирует рис. 12.3.3. Если тре-

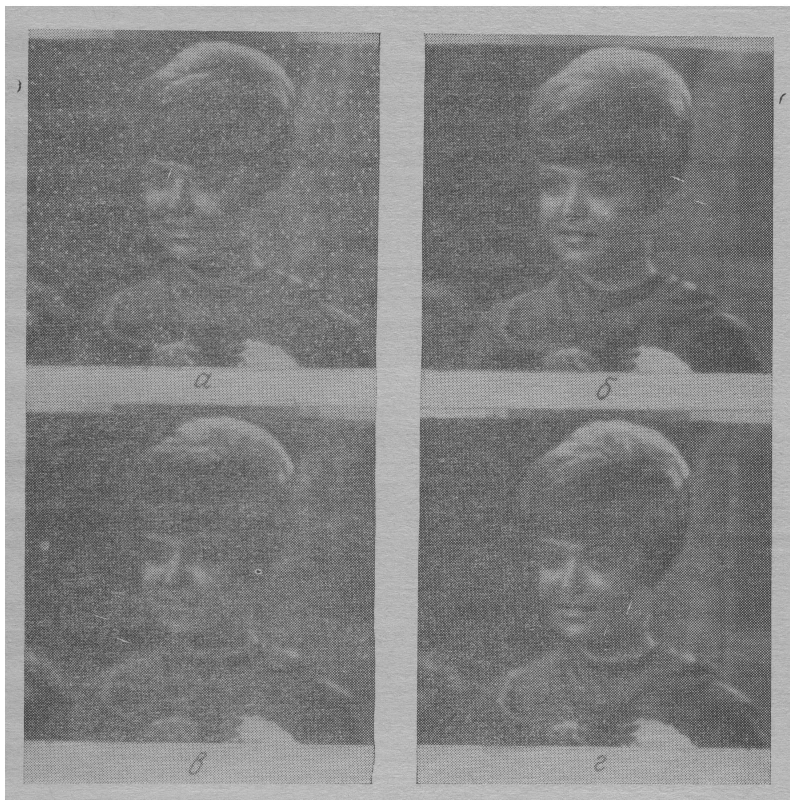


Рис. 12.3.3. Образцы изображений, подвергнутых низкочастотной фильтрации. *а* — зашумленное исходное изображение; *б* — действие маски 1, *в* — действие маски 2; *г* — действие маски 3

буемое подавление шума сопряжено с использованием массивов большого размера, целесообразно выполнять свертку косвенным образом, применяя преобразование Фурье (см. гл. 11), так как обычно это дает выигрыш в объеме вычислений.

12.4. ПОДЧЕРКИВАНИЕ ГРАНИЦ

Психофизические эксперименты показывают, что фотографическое или телевизионное изображение с подчеркнутыми границами часто оказывается субъективно более приятным, чем фотометрически совершенная репродукция. Метод подчеркивания границ можно реализовать несколькими способами.

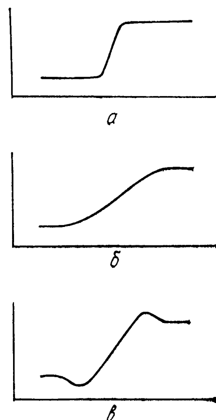


Рис. 12.4.1. Осциллограммы видеосигналов в системе улучшения изображений методом нерезкого маскирования.

а — высокая разрешающая способность; *б* — низкая разрешающая способность; *в* — эффект нерезкого маскирования.

В системах электронного сканирования изображений получаемый видеосигнал можно пропустить через электрический фильтр верхних частот. Другой способ обработки сканируемых изображений заключается в использовании нерезкого маскирования [12]. При этом изображение как бы сканируется двумя перекрывающимися апертурами, одна из которых соответствует нормальному разрешению, а другая — пониженному. В результате получают соответственно массив нормального изображения $F(j, k)$ и массив нечеткого изображения $F_L(j, k)$. Затем электронным способом формируют массив маскированного изображения

$$F_{\text{М}}(j, k) = cF(j, k) - (1 - c) F_L(j, k), \qquad (12.4.1)$$

где c — коэффициент пропорциональности. Обычно значения c находятся в пределах от $^3/_5$ до $^5/_6$, т. е. отношение составляющих нормальной и пониженной четкости изменяется от 1,5 до 5. На рис. 12.4.1 показаны типичные осциллограммы видеосигналов при сканировании резкой границы. Сигнал, полученный в результате маскирования, имеет два выброса, отсутствующие в исходном сигнале (высокого разрешения). Длительность фронта стала несколько больше. Субъективная резкость маскированного изображения повышается.

Подчеркивание границ можно также осуществить, выполняя дискретную фильтрацию согласно соотношению (12.3.1) с использованием высокочастотного импульсного отклика H [13—15].

Ниже представлены три типичные маски для выполнения высокочастотной фильтрации:

маска 1

$$\mathbf{H} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \qquad (12.4.2)$$

маска 2

$$\mathbf{H}=\begin{bmatrix}-1 & -1 & -1\\ -1 & 9 & -1\\ -1 & -1 & -1\end{bmatrix}, \qquad (12.4.3)$$

маска 3

$$\mathbf{H}=\begin{bmatrix}1 & -2 & 1\\ -2 & 5 & -2\\ 1 & -2 & 1\end{bmatrix}. \qquad (12.4.4)$$

Эти маски отличаются тем, что сумма их элементов равна единице. Рис. 12.4.2 иллюстрирует два способа подчеркивания границ: нерезкое маскирование и применение масок высокочастотной фильтрации.

Еще одним способом подчеркивания границ является так называемое статистическое дифференцирование [7, 16, стр. 100]. Значение яркости каждого элемента делится на статистическую оценку среднеквадратического отклонения $\sigma(j, k)$:

$$G(j, k) = F(j, k)/\sigma(j, k). \qquad (12.4.5)$$

Среднеквадратическое отклонение

$$\sigma^2(j, k) = \underbrace{\sum_j \sum_k}_{j,\, k \in N(j,\, k)} [F(j, k) - \overline{F}(j, k)]^2 \qquad (12.4.6)$$

вычисляется в некоторой окрестности $N(j, k)$ элемента с координатами (j, k). Функция $\overline{F}(j, k)$ — среднее значение яркости исходного изображения в точке с координатами (i, k), приближенно определяемое путем сглаживания изображения с помощью оператора низкочастотной фильтрации согласно формуле (12.3.1). Улучшенное изображение, представленное массивом $G(i, k)$, отличается от исходного изображения тем, что его яркость выше на границах, элементы которых не похожи на соседние элементы, и ниже на всех остальных участках. Следует отметить, что подчеркивание полезных границ сопровождается возрастанием шумовых составляющих.

Уоллис [17] обобщил метод статистического дифференцирования. Предложенный им оператор дает улучшенное изображение, имеющее требуемые моменты первого и второго порядка. Оператор определяется выражением вида

$$G(j, k) = [F(j, k) - \overline{F}(j, k)] \left[\frac{A\sigma_d}{A\sigma(j, k) + \sigma_d}\right] + \\ + [\alpha m_d + (1-\alpha)\overline{F}(j, k)]. \qquad (12.4.7)$$

Здесь m_d и σ_d — желаемые среднее значение и среднеквадратическое отклонение, A — коэффициент усиления, введенный для

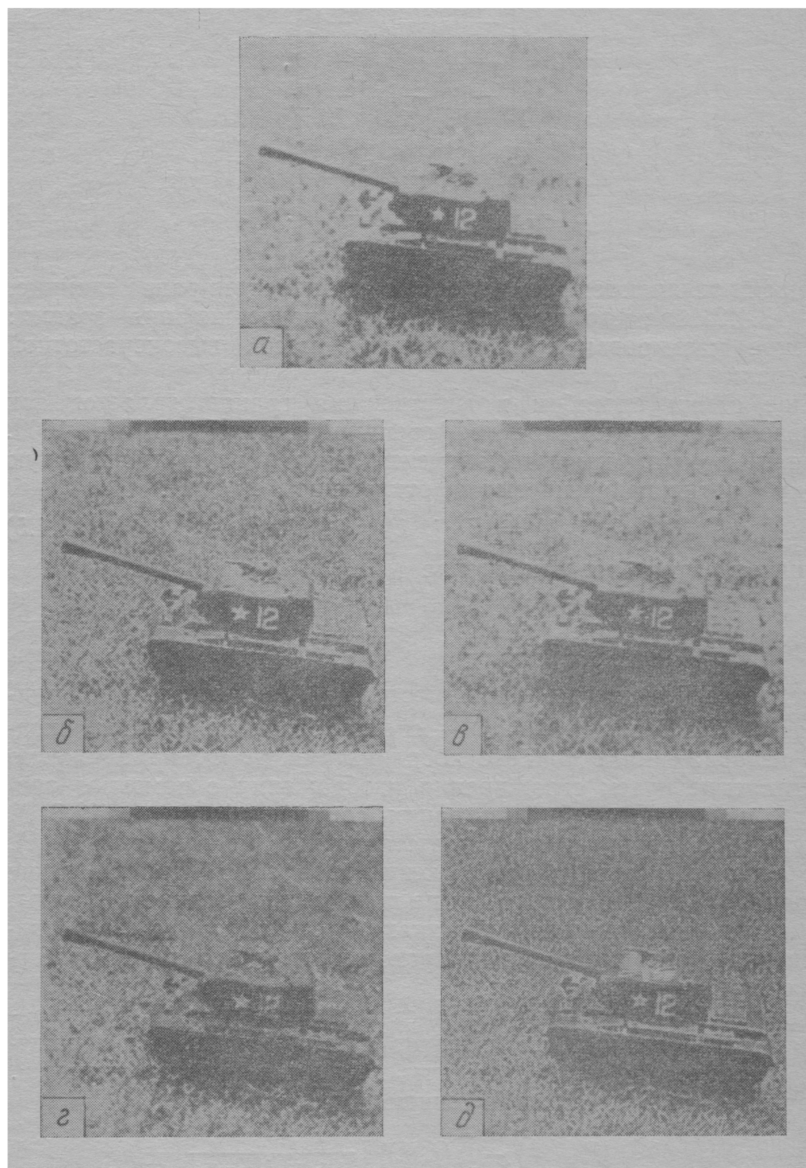


Рис. 12.4.2. Образцы изображений с подчеркнутыми границами при использовании нерезкого маскирования и масок высокочастотной фильтрации.

а — исходное изображение, *б* — нерезкое маскирование при $c = {}^{3}/_{4}$; *в* — действие маски 1; *г* — действие маски 2; *д* — действие маски 3.

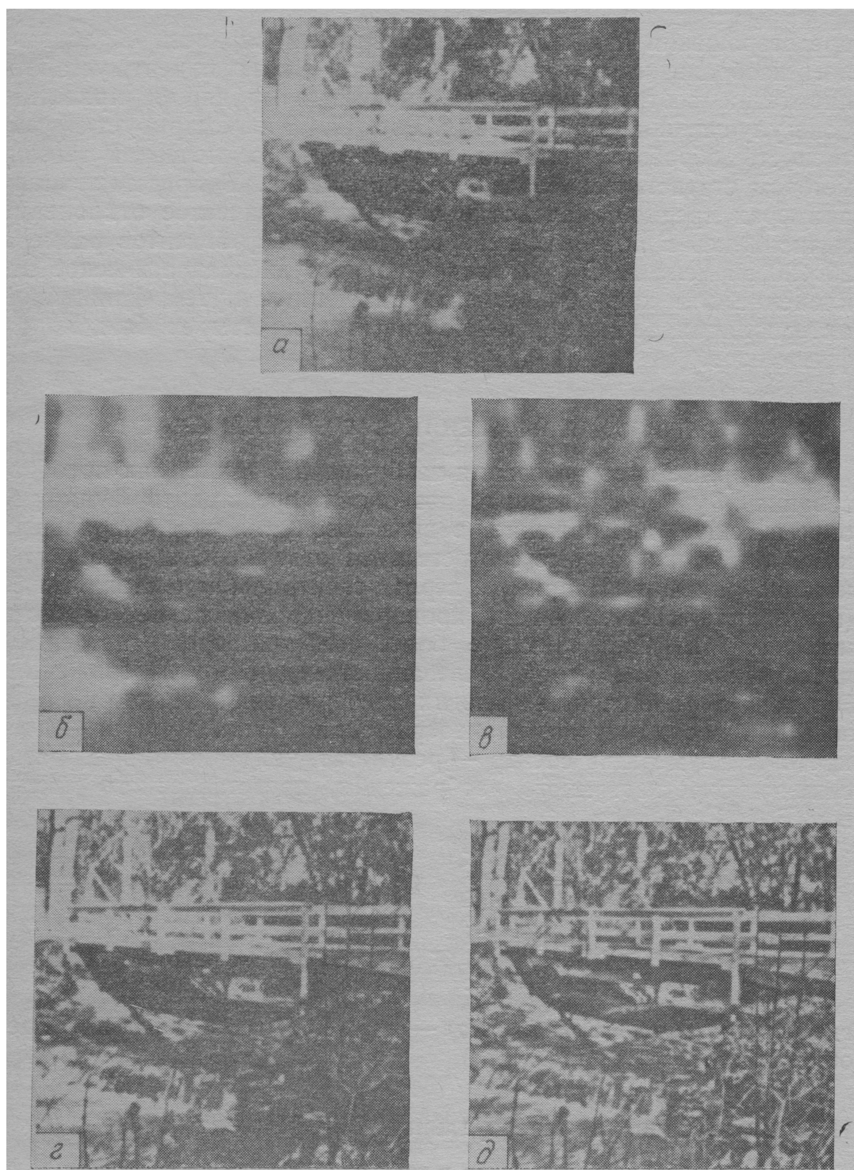


Рис. 12.4.3. Образцы изображений, обработанных методом статистического дифференцирования [17]. Фотографии предоставлены Р. Уоллисом из фирмы Stanford Technology (Саннивэйл, шт. Калифорния).

а — исходное изображение, *б* — результат усреднения по окрестности, *в* — результат вычисления среднеквадратического отклонения по окрестности, *г* — улучшенное изображение при $\alpha = 0{,}8$, $m_d = 128$, $\sigma_d = 85$, $A = 6$; *д* — улучшенное изображение при $\alpha = 0{,}1$, $m_d = 128$, $\sigma_d = 85$, $A = 6$

предотвращения излишне больших значений яркости обработанного изображения при малом $\sigma(j, k)$, α — коэффициент, устанавливающий соотношение между яркостью границ и яркостью фона улучшенного изображения. Рис. 12.4.3 иллюстрирует эффективность обработки сюжета с глубокими солнечными тенями методом статистического дифференцирования. В представленных примерах среднее значение и среднеквадратическое отклонение вычислялись в пределах непересекающихся фрагментов размера 20×20 элементов. Среднее значение и среднеквадратическое отклонение для каждого элемента находились путем билинейной интерполяции значений, вычисленных для четырех ближайших фрагментов.

12.5. ОБРАБОТКА ИЗОБРАЖЕНИЙ С ПРЕОБРАЗОВАНИЕМ

Унитарные преобразования изображений, такие, как преобразования Фурье и Адамара, позволяют представить функцию, описывающую изображение, в виде совокупности спектральных коэффициентов, которые соответствуют отдельным характеристикам изображения. Например, первая спектральная составляющая (постоянная составляющая) пропорциональна средней яркости изображения. Составляющие более высокой пространственной частоты являются мерой «изрезанности» данного изображения. Это характерное свойство многих преобразований можно использовать в целях улучшения изображений. Пусть $\mathcal{F}(u, v)$ — дискретный двумерный спектр — результат унитарного преобразования дискретизованного изображения, представленного массивом $F(i, k)$:

$$\mathcal{F}(u, v)=\sum_{j=0}^{N-1}\sum_{k=0}^{N-1} F(j, k)\, A(j, k; u, v), \qquad (12.5.1а)$$

$$F(j, k)=\sum_{u=0}^{N-1}\sum_{v=0}^{N-1} \mathcal{F}(u, v)\, B(j, k; u, v), \qquad (12.5.1б)$$

где $A(j, k; u, v)$ и $B(f, k; u, v)$ — ядра прямого и обратного преобразований.

Извлечение корня из спектральных коэффициентов [3]

Эта процедура заключается в том, что каждый из спектральных коэффициентов возводится в степень, причем знак или фаза коэффициентов сохраняются. Видоизмененный коэффициент определяется как

$$\tilde{\mathcal{F}}(u, v)=\frac{\mathcal{F}(u, v)}{|\mathcal{F}(u, v)|}|\mathcal{F}(u, v)|^{\alpha}=\mathcal{F}(u, v)\,|\mathcal{F}(u, v)|^{\alpha-1}. \qquad (12.5.2)$$

Для спектра Фурье, имеющего структуру

$$\mathcal{F}(u, v) = \mathcal{M}(u, v) \exp\{i\Phi(u, v)\}, \tag{12.5.3}$$

где $\mathcal{M}(u, v)$ и $\Phi(u, v)$ — соответственно амплитудная и фазовая составляющие данного коэффициента, видоизмененный коэффициент запишется как

$$\tilde{\mathcal{F}}(u, v) = [\mathcal{M}(u, v)]^{\alpha} \exp\{i\Phi(u, v)\}. \tag{12.5.4}$$

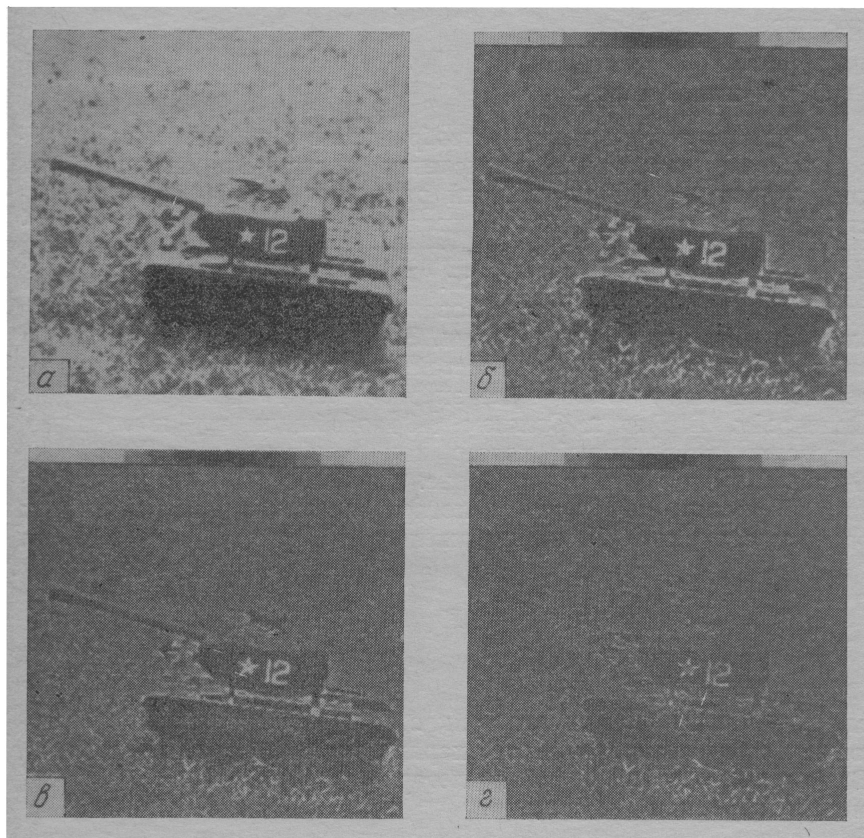


Рис. 12.5.1. Образцы изображений, обработанных методом извлечения корня из коэффициентов Фурье.

а — исходное изображение, $\alpha = 1$; *б* — изображение после извлечения квадратного корня, $\alpha = 1/2$; *в* — изображение после извлечения кубического корня, $\alpha = 1/3$; *г* — изображение при сохранении только фазового множителя спектра, $\alpha = 0$.

В предельном случае, когда $\alpha = 0$, видоизмененный спектральный коэффициент есть просто фазовый множитель исходного коэффициента. Если показатель степени α меньше единицы, то операция

(12.5.2) — извлечение корня из коэффициентов — обычно дает следующий результат: большие коэффициенты преобразования уменьшаются, малые — увеличиваются. Такое перераспределение энергии в частотной плоскости часто приводит к более эффективному использованию динамического диапазона изображения, а также улучшает субъективное восприятие воспроизводимого изображения. На рис. 12.5.1 приведены образцы изображений, обработанных методом извлечения корня из спектральных коэффициентов.

Обобщенный кепстр

Другой метод улучшения изображений с нелинейным преобразованием основан на вычислении логарифма спектральных коэффициентов. Количественный результат такой обработки состоит в том, что динамический диапазон спектральных коэффициентов обычно сужается, а динамический диапазон восстановленного изображения расширяется. Богерт, Хили и Тьюки [18] предложили подобное преобразование как средство обнаружения эхо-сигналов при обработке речи. Это преобразование, результат которого назван кепстром [1)], состоит в том, что вычисляется энергетический спектр логарифма энергетического спектра сигнала. Оппенгейм, Шафер и Стокхэм [19] обобщили эту концепцию, определив комплексный кепстр как результат преобразования Фурье логарифма спектра Фурье. Для произвольного преобразования кепстр обобщенного логарифмического преобразования определяется следующим выражением [20]:

$$\tilde{F}(j, k) = K \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} [\ln \{a + b\,|\mathcal{F}(u, v)|\}] \frac{\mathcal{F}(u, v)}{|\mathcal{F}(u, v)|} B(j, k; u, v), \tag{12.5.5}$$

где K — масштабная постоянная, a и b — постоянные, определяющие форму логарифмической характеристики передачи уровней. Выражение для обобщенного логарифмического преобразования можно переписать в виде

$$F(j, k) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \mathcal{G}(u, v) \mathcal{F}(u, v) B(j, k; u, v), \tag{12.5.6}$$

где

$$\mathcal{G}(u, v) = \frac{K \ln \{a + b\,|\mathcal{F}(u, v)|\}}{|\mathcal{F}(u, v)|} \tag{12.5.7}$$

---

1) Слово «кепстр» образовано перестановкой первых четырех букв в слове «спектр».

представляет собой частотную характеристику нелинейного фильтра, зависящую от вида исходного изображения, подлежащего фильтрации.

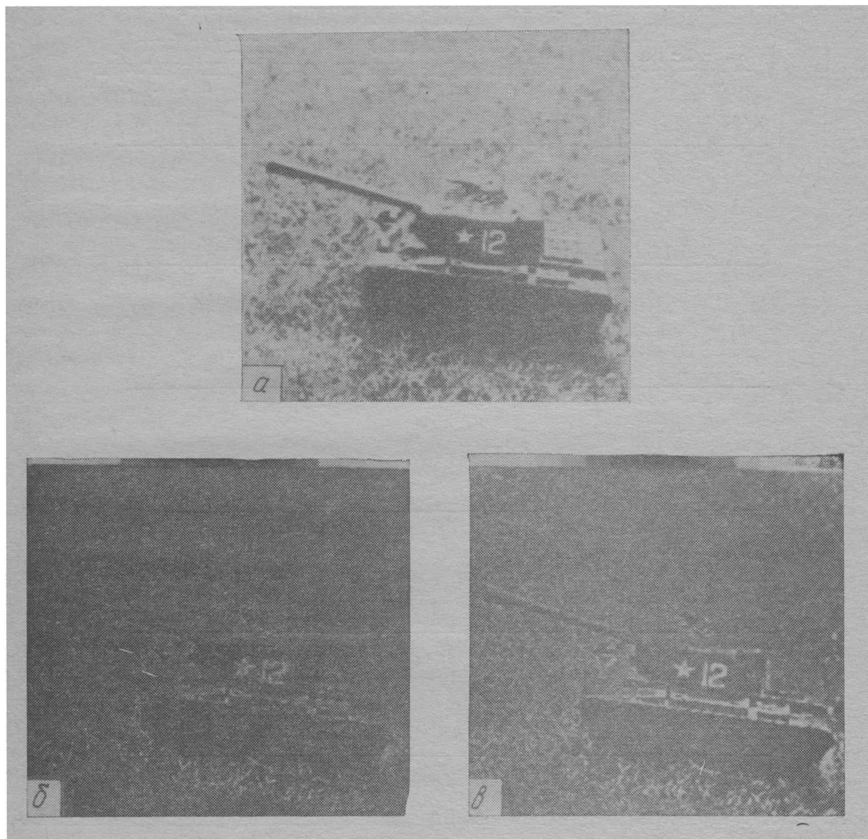


Рис. 12.5.2. Образцы изображений, обработанных путем вычисления кепстров Фурье и Адамара.

а — исходное изображение, *б* — изображение, соответствующее кепстру Фурье; *в* — изображение, соответствующее кепстру Адамара.

Изображения естественного происхождения обычно имеют большие низкочастотные и малые высокочастотные пространственные составляющие. Поэтому процедура логарифмического преобразования автоматически обеспечивает определенное подчеркивание границ изображения. На рис. 12.5.2 показаны образцы изображений, обработанных путем вычисления кепстров Фурье и Адамара.

12.6. МЕДИАННЫЙ ФИЛЬТР

Медианная фильтрация — метод нелинейной обработки сигналов, разработанный Тьюки [21]. Этот метод оказывается полезным при подавлении шума на изображении. Одномерный медиан-

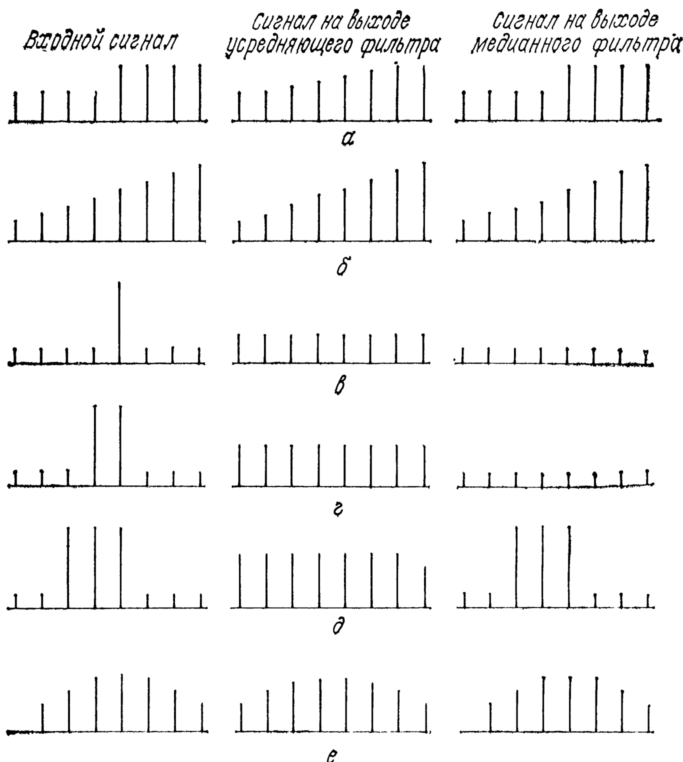


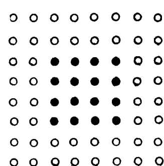
Рис. 12.6.1. Примеры медианной фильтрации простейших дискретных сигналов, $L = 5$.

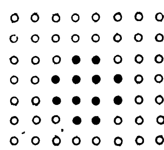
а — ступенчатый переход; *б* — пилообразный переход; *в* — одиночный импульс; *г* — сдвоенный импульс; *д* — строенный импульс; *е* — треугольный сигнал.

ный фильтр представляет собой скользящее окно, охватывающее нечетное число элементов изображения. Центральный элемент заменяется медианой всех элементов изображения в окне. Медианой дискретной последовательности $a_1, a_2, \ldots, a_N$ для нечетного N является тот ее элемент, для которого существуют $(N - 1)/2$ элементов, меньших или равных ему по величине, и $(N - 1)/2$ элементов, больших или равных ему по величине. Пусть в окно попали элементы изображения с уровнями 80, 90, 200, 110 и 120; в этом случае центральный элемент следует заменить значением

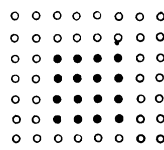
110, которое является медианой упорядоченной последовательности 80, 90, 110, 120, 200. Если в этом примере значение 200 является шумовым выбросом в монотонно возрастающей последовательности, то медианная фильтрация обеспечит существенное улучшение. Напротив, если значение 200 соответствует полезному импульсу сигнала (при использовании широкополосных датчиков), то обработка приведет к потере четкости воспроизводимого изображения. Таким образом, медианный фильтр в одних случаях обеспечивает подавление шума, в других — вызывает нежелательное подавление сигнала.

На рис. 12.6.1 показано воздействие медианного и усредняющего (сглаживающего) фильтров с пятиэлементным окном на ступенчатый, пилообразный, импульсный и треугольный дискретные сигналы. Из этих диаграмм видно, что медианный фильтр не влияет на ступенчатые или пилообразные функции, что обычно является желательным свойством. Однако этот фильтр подавляет импульсные сигналы, длительность которых составляет менее половины ширины окна. Фильтр также вызывает уплощение вершины треугольной функции.









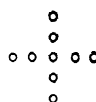


Рис. 12.6.2. Примеры двумерной медианной фильтрации.

Возможности анализа действия медианного фильтра ограничены. Можно показать, что медиана произведения постоянной K и последовательности $f(j)$ равна

$$\operatorname{med}\{Kf(j)\} = K \operatorname{med}\{f(j)\}. \tag{12.6.1}$$

Кроме того,

$$\operatorname{med}\{K + f(j)\} = K + \operatorname{med}\{f(j)\}. \tag{12.6.2}$$

Однако медиана суммы двух произвольных последовательностей $f(j)$ и $g(j)$ не равна сумме их медиан:

$$\operatorname{med}\{f(j) + g(j)\} \neq \operatorname{med}\{f(j)\} + \operatorname{med}\{g(j)\}. \tag{12.6.3}$$

Это неравенство можно проверить на примере последовательностей 80, 90, 100, 110, 120 и 80, 90, 100, 90, 80.

Возможны различные стратегии применения медианного фильтра для подавления шумов. Одна из них рекомендует начинать с медианного фильтра, окно которого охватывает три элемента

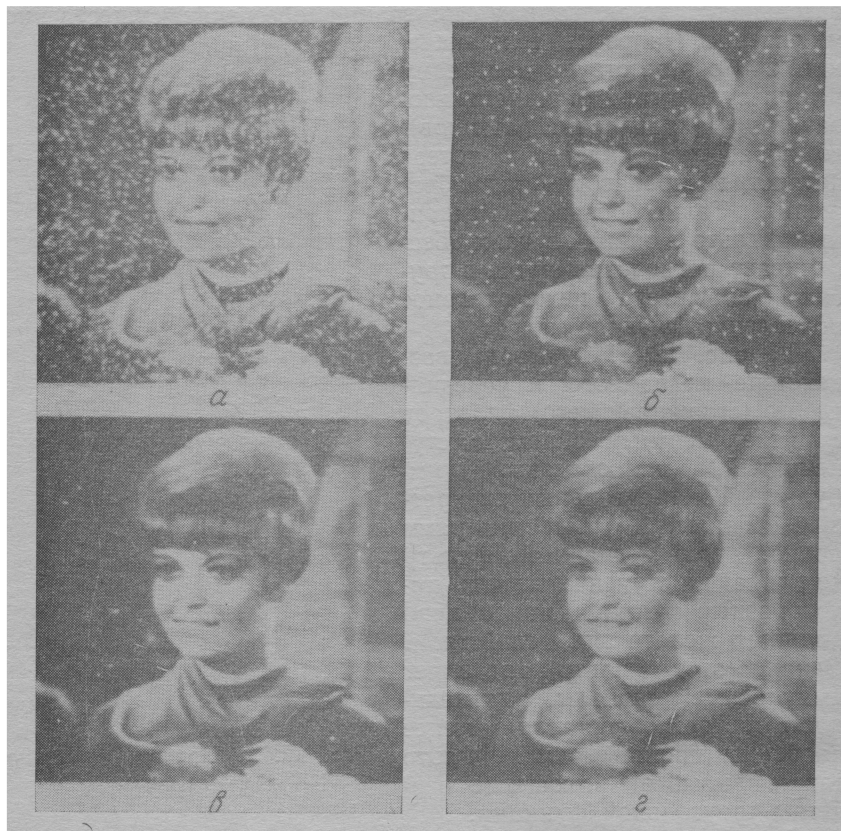


Рис. 12.6.3. Образцы изображений, обработанных одномерным медианным фильтром с целью подавления импульсных помех.

а — исходное изображение с импульсными помехами (15 искаженных элементов в каждой строке), *б* — результат медианной фильтрации при $L = 3$; *в* — результат медианной фильтрации при $L = 5$, *г* — результат медианной фильтрации при $L = 7$.

изображения. Если ослабление сигнала незначительно, окно фильтра расширяют до пяти элементов. Так поступают до тех пор, пока медианная фильтрация начинает приносить больше вреда, чем пользы. Другая возможность состоит в осуществлении каскадной медианной фильтрации сигнала с использованием фиксированной или изменяемой ширины окна. В общем случае те области, которые остаются без изменения после однократной обработки фильтром, не меняются и после повторной обработки. Области,

в которых длительность импульсных сигналов составляет менее половины ширины окна, будут подвергаться изменениям после каждого цикла обработки.



Рис. 12.6.4. Образцы изображений, обработанных одномерным медианным фильтром с целью подавления гауссова шума.

а — исходное изображение с гауссовым шумом $\sigma_n = 25$; *б* — результат медианной фильтрации при $L = 3$; *в* — результат медианной фильтрации при $L = 5$; *г* — результат медианной фильтрации при $L = 7$.

Концепцию медианного фильтра легко обобщить на два измерения, применяя двумерное окно желаемой формы, например прямоугольное или близкое к круговому. Очевидно, что двумерный медианный фильтр с окном размера $L \times L$ обеспечивает более эффективное подавление шума, чем последовательно примененные горизонтальный и вертикальный одномерные медианные фильтры с окном размера $L \times 1$; двумерная обработка, однако, приводит к более существенному ослаблению сигнала. На рис. 12.6.2 пока-

зано влияние двумерной медианной фильтрации на пространственный импульсный сигнал. Использовались фильтры с окнами двух типов: квадратным размера 3×3 и крестообразным размера 5×5. Как видно, медианный фильтр с квадратным окном разрушил углы изображенного квадрата, а фильтр с крестообразным окном оставил его без изменения.

На рис. 12.6.3 и 12.6.4 представлены образцы изображений, подвергнутых медианной фильтрации с целью подавления шумов. Как видно, медианный фильтр более эффективно подавляет разрозненные импульсные помехи, чем гладкие шумы. Медианную фильтрацию изображений в целях подавления шумов следует считать эвристическим методом. Ее нельзя применять вслепую. Напротив, следует проверять получаемые результаты, чтобы убедиться в целесообразности медианной фильтрации.

12.7. ЛОЖНЫЕ ЦВЕТА

Ложные цвета образуются при поэлементном линейном или нелинейном преобразовании координат цвета исходного цветного изображения или набора компонент спектрозонального изображения, в результате которого получаются координаты воспроизводимого цвета [22, 23]. Это преобразование применяется для того, чтобы получить изображение, объекты которого имеют измененные или ложные цвета, отличающиеся от ожидаемых. Например, голубое небо в естественном сюжете может стать красным, а зеленая трава — синей. Одна из возможных целей такого цветового преобразования — поместить обыкновенные объекты в мир необычных цветов, чтобы наблюдатель следил за ними более внимательно, чем в случае их нормальной окраски. Медленное изменение преобразования позволяет предотвратить адаптацию наблюдателя к новому цветовому миру. Польза такого применения ложных цветов связана со сложными психологическими явлениями.

Ложные цвета могут применяться также для лучшего использования возможностей зрительной системы человека. Как известно, яркостная чувствительность палочек и колбочек сетчатки максимальна в зеленой области видимого спектра. Таким образом, если объект красного цвета перекрасить в ложный зеленый цвет, его будет легче обнаружить. Другое психофизическое свойство цветового зрения, которым можно воспользоваться, заключается в повышенной контрастной чувствительности глаза в синем свете. При определенных обстоятельствах целесообразно преобразовывать естественные цвета объектов с мелкими деталями в градации синего цвета.

Третье применение ложных цветов связано с воспроизведением в естественных цветах спектрозональных изображений. При этом

некоторые компоненты спектрозонального изображения могут формироваться датчиками, спектральные характеристики которых находятся вне видимой области спектра, например в инфракрасной или ультрафиолетовой области.

При образовании ложных цветов красная, зеленая и синяя координаты воспроизводимых цветов связаны с координатами исходных цветов или значениями F_i компонент спектрозональных изображений следующими выражениями:

$$R_D = \mathcal{O}_R\{F_1, F_2, \ldots\}, \quad (12.7.1а)$$

$$G_D = \mathcal{O}_G\{F_1, F_2, \ldots\}, \quad (12.7.1б)$$

$$B_D = \mathcal{O}_B\{F_1, F_2, \ldots\}, \quad (12.7.1в)$$

где $\mathcal{O}_R(\cdot)$, $\mathcal{O}_G(\cdot)$, $\mathcal{O}_B(\cdot)$ — функциональные операторы общего вида. Рассмотрим простой пример, когда исходные красная, зеленая и синяя координаты цвета ($R_S = F_1$, $G_S = F_2$, $B_S = F_3$) меняются местами в соответствии с соотношением

$$\begin{bmatrix} R_D \\ G_D \\ B_D \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} R_S \\ G_S \\ B_S \end{bmatrix}. \quad (12.7.2)$$

Зеленые объекты исходного изображения будут воспроизводиться как красные, синие — как зеленые, а красные — как синие. Общее выражение для линейного преобразования в ложные цвета естественных цветных изображений можно определить как

$$\begin{bmatrix} R_D \\ G_D \\ B_D \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} R_S \\ G_S \\ B_S \end{bmatrix}. \quad (12.7.3)$$

Следует заметить, что это соотношение описывает линейное преобразование координатной системы основных цветов датчика исходного изображения в координатную систему некоторых новых основных цветов.

12.8. ПСЕВДОЦВЕТА

Псевдоцветовое отображение является еще одним методом повышения обнаружимости изображенных объектов человеком-наблюдателем. Однако при использовании псевдоцветов исходное изображение не является цветным, а представляется двумерным массивом величин, отображаемых в цветовое пространство [24—

26]. Псевдоцветовое отображение элемента $F(j, k)$ исходного массива в общем виде определяется как

$$R(j, k) = \mathcal{O}_R\{F(j, k)\}, \quad (12.8.1а)$$

$$G(j, k) = \mathcal{O}_G\{F(j, k)\}, \quad (12.8.1б)$$

$$B(j, k) = \mathcal{O}_B\{F(j, k)\}, \quad (12.8.1в)$$

где $R(j, k)$, $G(j, k)$, $B(j, k)$ — координаты цвета воспроизводимых цветов, $\mathcal{O}_R(\cdot)$, $\mathcal{O}_G(\cdot)$, $\mathcal{O}_B(\cdot)$ — линейные или нелинейные функциональные операторы. Это отображение параметрически определяет некоторую траекторию в трехмерном цветовом пространстве, причем параметром является значение элемента $F(j, k)$. На рис. 12.8.1 показаны цветовое RGB-пространство и два псевдоцветовых отображения, траектории которых начинаются в точке черного и кончаются в точке белого. Отображение A представляется ахроматической траекторией, проходящей через все градации серого; оно является обычным представлением черно-белого изображения. Отображение B представляется некоторой спиральной траекторией в цветовом пространстве.

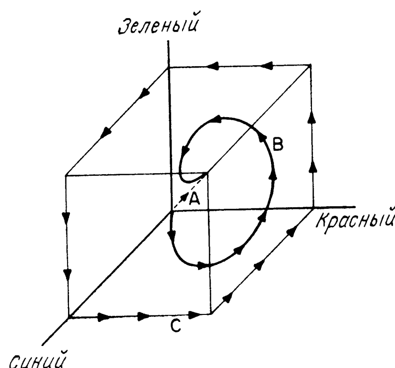


Рис. 12.8.1. Псевдоцветовые отображения, траектории которых начинаются в точке черного и кончаются в точке белого, и отображение, соответствующее обходу по периметру RGB-треугольника.

Другой класс псевдоцветовых отображений составляют отображения, не охватывающие градаций серого. Примером может служить отображение C, траектория которого проходит вдоль ребер цветового RGB-куба. Это отображение соответствует продвижению по периметру треугольника воспроизводимых цветов, который построен на равноконтрастном цветовом графике, показанном на рис. 12.8.2. На этом рисунке указаны яркости красного, зеленого, синего, голубого, пурпурного и желтого цветов, координаты которых лежат на сторонах треугольника воспроизводимых цветов. Как видно, яркость псевдоцветов изменяется в интервале между минимальным значением 0,114 (яркость синего) и максимальным значением 0,886 (яркость желтого). Максимальное значение яркости, равное единице, имеет белый цвет. В некоторых применениях желательна такая яркость всех воспроизводимых цветов, чтобы различение псевдоцветов по всей шкале проводилось только на основе цветового тона и насыщенности. Геометрические места точек равной яркости показаны на рис. 12.8.2.

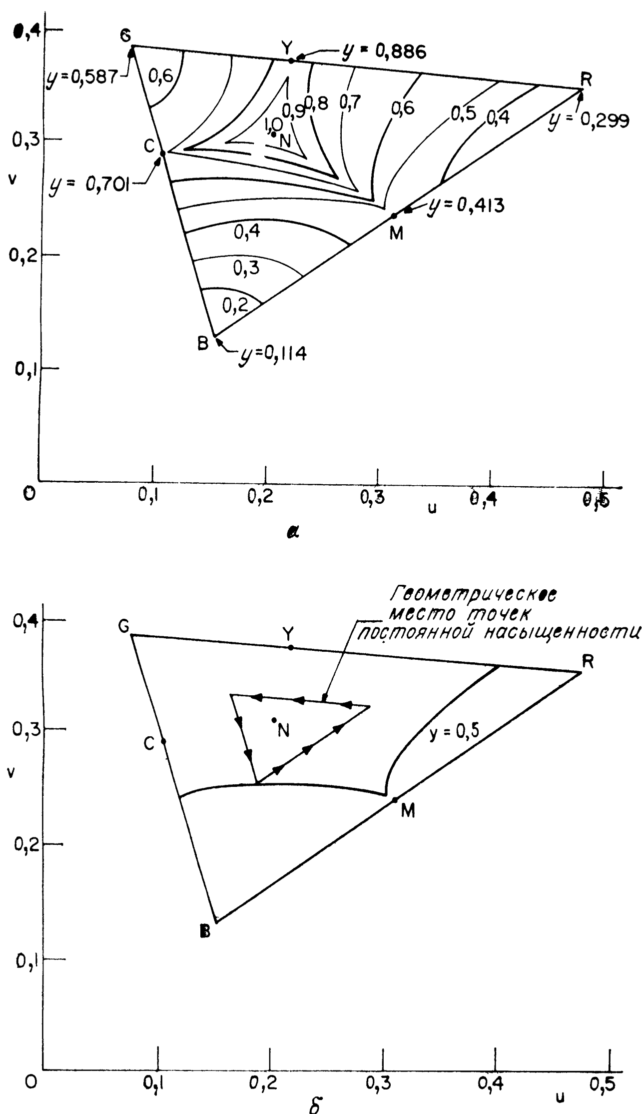


Рис. 12.8.2. Геометрические места точек постоянной яркости для цветов системы НТСЦ.

а — геометрические места точек максимально допустимых яркостей; *б* — геометрическое место точек для яркости, равной 0,5.

На этом рисунке также представлены границы цветностей воспроизводимых цветов равной яркости. Например, обход по периметру *RGB*-треугольника возможен только при ограничении максимальной яркости любого из цветов значением 0,114, что соответствует яркости синего. Для яркости 0,2 обход по периметру *RGB*-треугольника также возможен, за исключением области вблизи насыщенного синего. Для больших уровней яркости гамма цветов равной яркости, пригодных для включения в шкалу псевдоцветов, становится существенно ограниченной. На рис. 12.8.2,*б* показано геометрическое место цветностей для яркости 0,5. В область, ограниченную этим геометрическим местом точек, вписан треугольник цветностей тех цветов, которые имеют одинаковую наибольшую насыщенность. Псевдоцвета с цветностями на сторонах этого треугольника будут отличаться только цветовым тоном.

При заданной линии псевдоцветов в цветовом пространстве необходимо выбрать масштабный коэффициент, связывающий значения $F(j, k)$ с расстоянием вдоль этой линии. На равноконтрастном цветовом графике равным приращениям расстояния соответствуют субъективно почти одинаковые изменения цветового ощущения. Это дает основание делить всю длину линии псевдоцветов на равные отрезки.

12.9. УЛУЧШЕНИЕ СПЕКТРОЗОНАЛЬНЫХ ИЗОБРАЖЕНИЙ

При обработке спектрозональных изображений часто используют различные процедуры для облегчения последующего визуального или машинного анализа [22, 27]. Некоторые из таких процедур предусматривают раздельную обработку спектрозональных компонент, например повышение контраста, видоизменение гистограмм и подчеркивание границ (эти процедуры описаны в предыдущих разделах). Однако существуют и другие процедуры, рассматриваемые в данном разделе, которые основаны на совместной обработке спектрозональных компонент.

Часто используют разности двух компонент спектрозональных изображений:

$$D_{m,n}(j, k) = F_m(j, k) - F_n(j, k), \qquad (12.9.1)$$

чтобы выявить различие между коэффициентами отражения в разных зонах спектра. Дополнительное преимущество этой операции состоит в подавлении неизвестного, но общего смещения яркости, которое может существовать. Другой простой, но эффективный способ улучшения спектрозональных изображений состоит в вычислении отношений отдельных компонент спектрозонального

изображения. Отношение m-й и n-й спектрозональных компонент равно по определению

$$R_{m,n}(j, k) = F_m(j, k)/F_n(j, k). \quad (12.9.2)$$

Предполагается, что диапазон изменения спектрозональных компонент выбран так, что они не принимают нулевых значений. Во многих спектрозональных системах компоненту $F_n(j, k)$ удается представить в виде произведения коэффициента отражения объекта $R_n(j, k)$ и функции освещенности $E(j, k)$; это произведение почти одинаково для всех компонент. Вычисление отношений компонент обеспечивает автоматическую нормировку, т. е. позволяет компенсировать изменение освещенности. Одна из трудностей, с которой сталкиваются при использовании метода отношений, заключается в подчеркивании ошибок квантования. Отношение $[F_m(j, k)/F_n(j, k) \pm \Delta(j, k)]$, в котором $\Delta(j, k)$ выражает неопределенность квантования, при малом $F_n(j, k)$ может изменяться в значительных пределах. Разброс существенно уменьшается при вычислении логарифма отношения, введенного в работе [26]:

$$L_{m,n}(j, k) = \log[R_{m,n}(j, k)] = \log[F_m(j, k)] - \log[F_n(j, k)]. \quad (12.9.3)$$

Имея N спектрозональных компонент, можно получить $[N(N-1)]/2$ различных разностей или отношений. Чтобы уменьшить число вариантов, разности и отношения часто вычисляют относительно усредненного изображения, описываемого массивом

$$A(j, k) = (1/N) \sum_{n=1}^{N} F_n(j, k). \quad (12.9.4)$$

Унитарные преобразования совокупности спектрозональных компонент представляют собой еще одно средство их улучшения. При наличии N спектрозональных компонент формируется вектор размера $N \times 1$:

$$\mathbf{x} = \begin{bmatrix} F_1(j, k) \\ F_2(j, k) \\ \vdots \\ F_N(j, k) \end{bmatrix} \quad (12.9.5)$$

для каждого элемента (j, k). Затем производится преобразование вида

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \quad (12.9.6)$$

где **A** — унитарная матрица размера $N \times N$. Часто используется преобразование, которое представляет собой разложение на главные компоненты, когда строки матрицы **A** составлены из

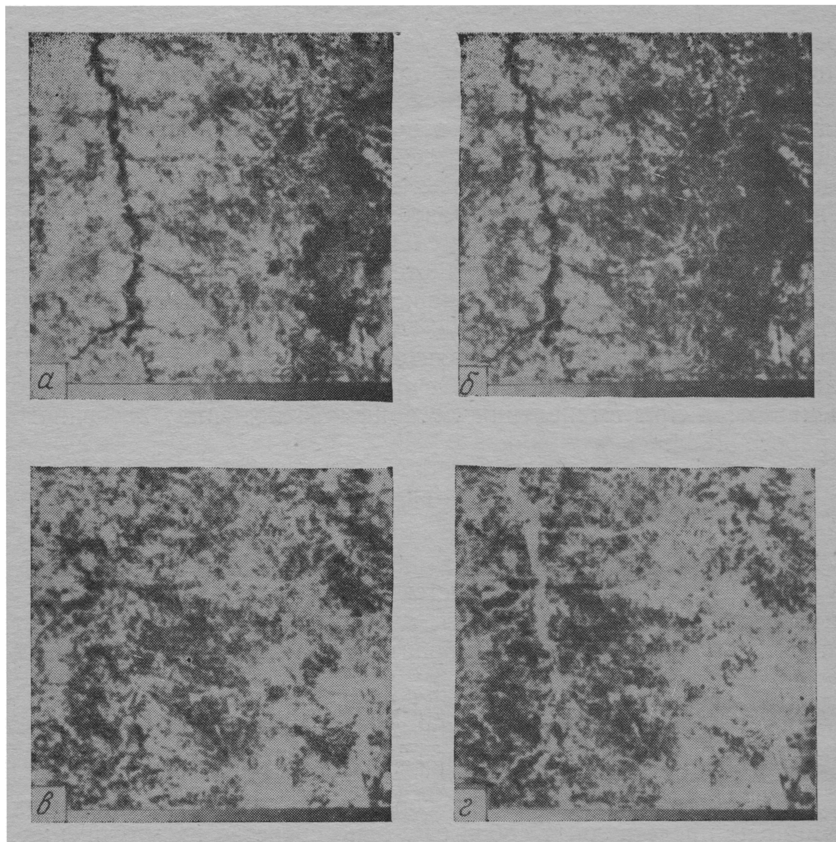


Рис. 12.9.1. Образцы спектрозональных компонент [27].

а — спектральный диапазон 4 (зеленая область); *б* — спектральный диапазон 5 (красная область); *в* — спектральный диапазон 6 (инфракрасная область 1); *г* — спектральный диапазон 7 (инфракрасная область 2).

собственных векторов ковариационной матрицы $\mathbf{K}_x$ спектрозональных компонент. Матрица **A** осуществляет диагонализацию ковариационной матрицы $\mathbf{K}_x$, такую, что ковариационная матрица преобразованных компонент

$$\mathbf{K}_y = \mathbf{A}\mathbf{K}_x\mathbf{A}^T = \mathbf{\Lambda} \tag{12.9.7}$$

становится диагональной матрицей $\mathbf{\Lambda}$, элементы которой представляют собой собственные значения матрицы $\mathbf{K}_x$, ранжированные

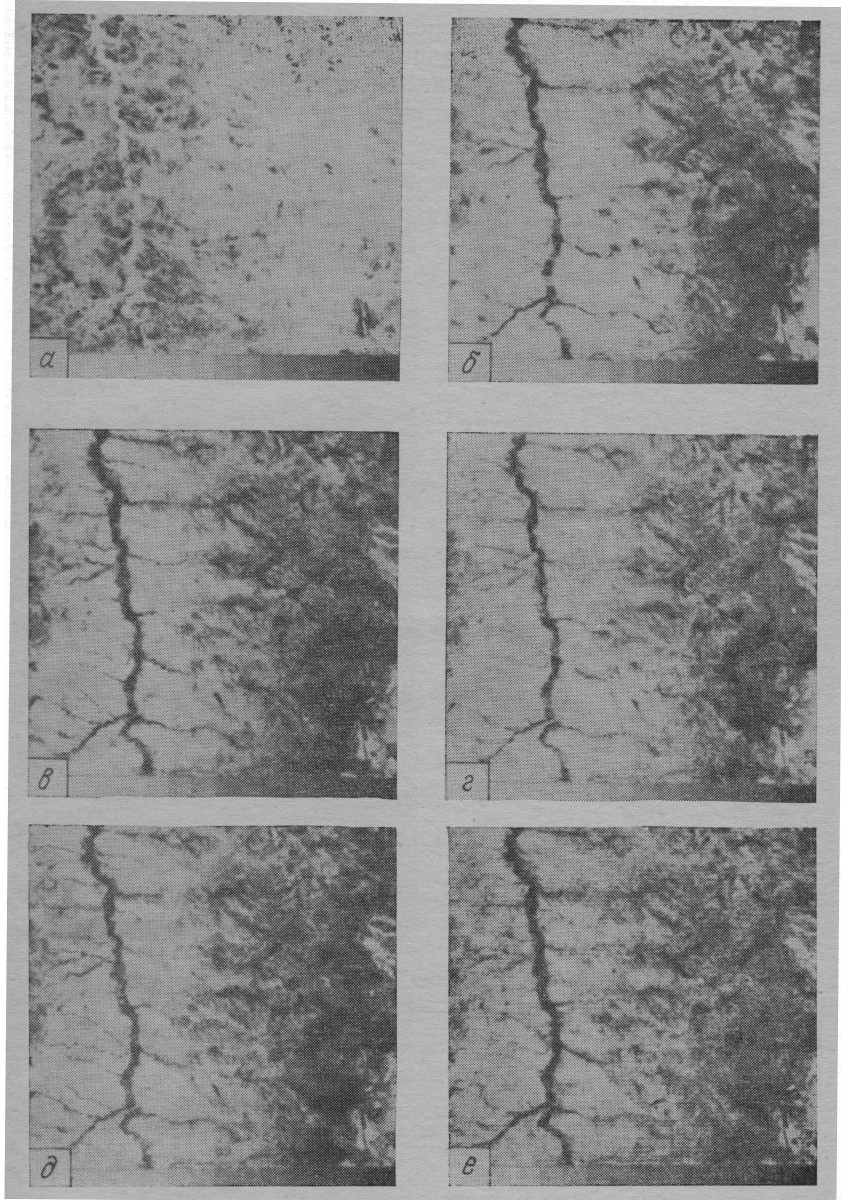


Рис. 12.9.2. Изображения, соответствующие логарифмам отношений компонент спектрального изображения, показанных на рис. 12.9.1.

а — изображение диапазона 4/изображение диапазона 5; *б* — изображение диапазона 4/изображение диапазона 7; *г* — изображение диапазона 5/изображение диапазона 6; *д* — изображение диапазона 5/изображение диапазона 7; *е* — изображение диапазона 6/изображение диапазона 7.

в порядке убывания. Поэтому в результате разложения на главные компоненты получают совокупность декоррелированных компонент с убывающими энергиями. Конечно, эта процедура требует знания ковариационной матрицы спектрозональных компо-

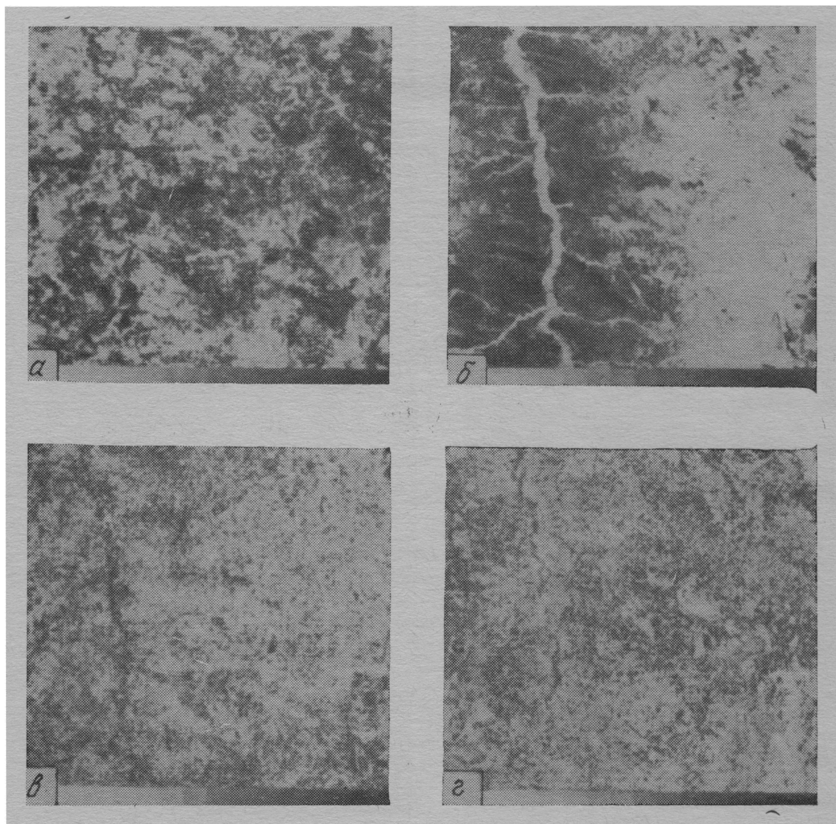


Рис. 12.9.3. Изображения, соответствующие главным компонентам спектрозональных изображений, которые представлены на рис. 12.9.1.

а — для собственного значения λ_1; *б* — для собственного значения λ_2; *в* — для собственного значения λ_3, *г* —для собственного значения λ_4

нент. Такую матрицу следует либо промоделировать, либо оценить. Если ковариационная матрица существенно нестационарна, метод разложения на главные компоненты применять затруднительно.

На рис. 12.9.1 показаны четыре компоненты спектрозонального изображения, а на рис. 12.9.2 — изображения, соответствующие логарифмам отношений этих компонент. Изображения, соответствующие главным компонентам, представлены на рис. 12.9.3.

ЛИТЕРАТУРА

1. Nathan R., Picture Enhancement for the Moon, Mars, and Man, in: Pictorial Pattern Recognition, Cheng G. C., Ed., Thompson, Washington D. C., 1968, pp. 239—266.
2. Billingsley F., Applications of Digital Image Processing, *Appl. Opt.*,**9**, 2, 289—299 (February 1970).
3. Andrews H. C., Tescher A. G., Kruger R. P., Image Processing by Digital Computer, *IEEE Spectrum*, **9**, 7, 20—32 (July 1972).
4. Hall E. L. et al., A Survey of Preprocessing and Feature Extraction Techniques for Radiographic Images, *IEEE Trans. Computers*, **C-20**, 9, 1032—1044 (September 1971).
5. Hall E. L., Almost Uniform Distribution for Computer Image Enhancement, *IEEE Trans. Computers*, **C-23**, 2, 207—208 (February 1974).
6. Frei W. (частное сообщение).
7. Ketcham D. J., Real Time Image Enhancement Technique, Proceedings SPIE/OSA Conference on Image Processing, Pacific Grove, California, Vol. 74, February 1976, pp. 120—125.
8. Dinneen G. P., Programming Pattern Recognition, Proceedings Western Joint Computer Conference, March 1955, pp. 94—100.
9. Graham R. E., Snow-Removal: A Noise-Stripping Process for Picture Signals, *IRE Trans. Inf. Theory*, IT-**8**, 1, 129—144 (February 1962).
10. Rosenfeld A., Park C. M., Strong J. P., Noise Cleaning in Digital Pictures, Proceedings EASCON Convention Record, October 1969, pp. 264—273.
11. Nathan R., Spatial Frequency Filtering, in: Picture Processing and Psychopictorics, Lipkin B. S., Rosenfeld A., Eds., Academic Press, New York, 1970, pp. 151—164.
12. Schreiber W. F., Wirephoto Quality Improvement by Unsharp Masking, *J. Pattern Recognition*, **2**, 171 (1970), Pergamon Press, London.
13. Roberts L. G., Machine Perception of Three-Dimensional Solids, in: Optical and Electro-Optical Information Processing, Tippett J. T. et al., Eds., MIT Press, Cambridge, Mass., 1965.
14. Prewitt J. M. S., Object Enhancement and Extraction in: Picture Processing and Psychopictorics, Lipkin B. S., Rosenfeld A., Eds., Academic Press, New York, 1970, pp. 75—150.
15. Arcese A., Mengert P. H., Trombini E. W., Image Detection Through Bipolar Correlation, *IEEE Trans. Inf. Theory*, **IT-16**, 5, 534—541 (September 1970).
16. Rosenfeld A., Picture Pıocessing by Computer, Academic Press, New York, 1969. [Имеется перевод: Розенфельд А., Распознавание и обработка изображений с помощью вычислительных машин. — М.: Мир, 1972.]
17. Wallis R. H., An Approach for the Space Variant Restoration and Enhancement of Images, Proceedings Symposium on Current Mathematical Problems in Image Science, Monterey, California, November 1976.
18. Bogert B. P., Healy M. J. R., Tukey J. W., The Frequency Analysis of Time Series for Echoes: Cepstrum and Shape Cracking, in: Proceedings of the Symposium held at Brown University, July 11—14, 1962, Rosenblatt M., Ed., Willey, New York, 1963, Ch. 15, pp. 209—243.
19. Oppenheim A. V., Schafer R. W., Stockham T. G., Jr., Nonlinear Filtering of Multiplied and Convolved Signals, *Proc. IEEE*, **56**, 8, 1264—1292 (August 1968). [Имеется перевод: Оппенгейм А. В., Шефер Р., Стокхэм Т., мл., Нелинейная фильтрация сигналов, представленных в виде произведения и свертки. — ТИИЭР, 1975, т. 56, № 8, с. 5—34].
20. Andrews H. C., Pratt W. K., Digital Image Transform Processing, Proceedings Applications of Walsh Functions, Washington D. C., 1970, pp. 183—194.
21. Tukey J. W., Exploratory Data Analysis, Addison-Wesley, Reading, Mass., 1971.

22. Goetz A. F. H. et al., Application of ERTS Images and Image Processing to Regional Geologic Problems and Geologic Mapping in Northern Arizona, Jet Propulsion Laboratory, Technical Report 32-1597, Pasadena, California, May 1975.
23. Fink W., Image Coloration as an Interpretation Aid, Proceedings SPIE/OSA Conference on Image Processing, Pacific Grove, California, Vol. 74, February 1976, pp. 209—215.
24. Gazley C., Reiber J. E., Stratton R. H., Computer Works a New Trick in Seeing Pseudo Color Processing, *Aeronautics and Astronautics*, **4**, 56 (April 1967).
25. Nichols L. W., Lamar J., Conversion of Infrared Images to Visible in Color, *Appl. Opt.*, 7, 9, 1757 (September 1968).
26. Kreins E. R., Allison L. J., Color Enhancement of Nimbus High Resolution Infrared Radiometer Data, *Appl. Opt.*, **9**, 3, 681 (March 1970).
27. Robinson G. S., Frei W., Final Research Report on Computer Processing of ERTS Images, University of Southern California, Image Processing Institute, Report USCIPI 640, September 1975.

Глава 13

МОДЕЛИ, ИСПОЛЬЗУЕМЫЕ ПРИ РЕСТАВРАЦИИ ИЗОБРАЖЕНИЙ

Реставрацию изображений можно рассматривать как процесс оценивания: некоторое изображение, полученное в результате наблюдения или измерения, подвергают преобразованию, чтобы найти сценку идеального изображения, которое наблюдалось бы на выходе гипотетической изображающей системы, не вносящей никаких искажений. В данной главе рассмотрены математические модели искажений, возникающих в изображающих системах общего вида. В последующих главах эти модели используются в качестве основы при разработке различных методов реставрации изображений.

13.1. ОБОБЩЕННЫЕ МОДЕЛИ, ИСПОЛЬЗУЕМЫЕ ПРИ РЕСТАВРАЦИИ ИЗОБРАЖЕНИЙ

Для эффективного проектирования любой системы цифровой реставрации изображений необходимо знать количественные оценки искажений, вносимых физической изображающей системой, преобразователем в цифровую форму и дисплеем. Процедура реставрации в основном сводится к моделированию, а затем к обращению искажающих преобразований. Следует подчеркнуть, что адекватное моделирование процесса, порождающего искажения, часто является ключом к эффективной реставрации.

Существуют два основных подхода к моделированию искажений: априорный и апостериорный. В первом случае измеряют отклики физической изображающей системы, преобразователя в цифровую форму и дисплея на произвольное изображение. Отклик одних систем можно получить детерминированным путем, отклик других поддается определению только в вероятностном смысле. В случае апостериорного моделирования модель искажений создают на основе измерений параметров конкретного искаженного изображения, подлежащего реставрации. По существу эти два подхода различаются только способами сбора информации, описывающей характер искажений.

На рис. 13.1.1 представлена обобщенная модель цифровой изображающей системы и системы реставрации изображений. На вход изображающей системы поступает световое поле $C(x, y, t, \lambda)$ — функция пространственных координат (x, y), времени t

и длины волны спектрального излучения λ. Это поле портится поэлементными и пространственными искажениями, а также детерминированными и случайными помехами. В числе возможных искажений можно назвать дифракционные искажения и аберрации оптической системы, нелинейные искажения фотоматериалов, искажения, обусловленные турбулентностью атмосферы, нерезкость (смаз), вызванную движением, геометрические искажения.

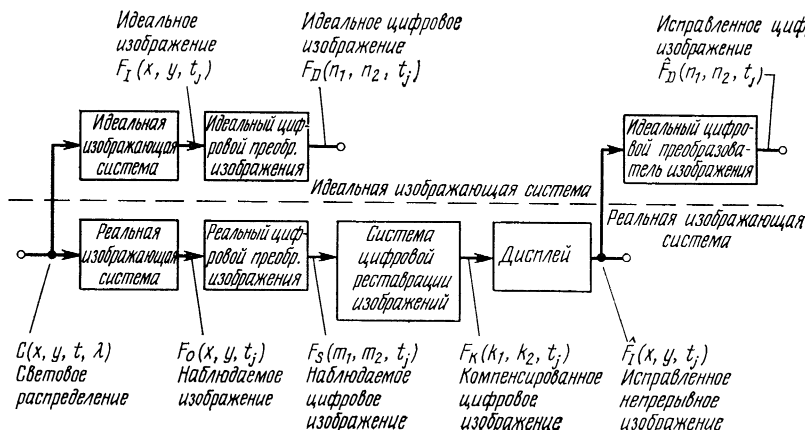


Рис. 13.1.1. Модель цифровой реставрации изображений.

Шумовые возмущения могут обусловливаться флуктуациями в электрических датчиках изображающих систем или зернистостью фотопленки. Изображающая система формирует в момент времени t совокупность выходных изображений, описываемых в общем виде выражением [1])

$$F_O^{(i)}(x, y, t) = \mathcal{O}_P\{C, x, y, t, \lambda; C(x, y, t, \lambda)\}, \quad (13.1.1)$$

где $\mathcal{O}_P\{\cdot\}$ — обобщенный оператор, являющийся функцией пространственных координат (x, y), времени t, длины волны λ и интенсивности светового поля C. В одноцветной изображающей системе формируется только одно выходное изображение. В цветной изображающей системе $F_O^{(i)}(\cdot)$ может относиться к изображениям, соответствующим координатам красного, зеленого и синего цвета ($i = 1, 2, 3$). В спектрозональной системе компоненты $F_O^{(i)}(\cdot)$ соответствуют различным зонам спектра.

1) Здесь принято следующее обозначение: переменные до точки с запятой показывают функциональную зависимость, т. е. наличие t означает, что $\mathcal{O}\{\cdot\}$ зависит от времени, а член после точки с запятой — функция, подвергаемая обработке.

Оператор обобщенной изображающей системы (13.1.1) может иметь различную структуру. Например, система может обладать или не обладать «памятью». Временной отклик представляет собой прямой результат взаимодействия наблюдаемых элементов изображения с одними и теми же координатами (x, y), но принадлежащих различным кадрам, разнесенным во времени. В рамках этого представления пространственная нерезкость, обусловленная апертурными искажениями электронного растрового сканера, не связывается с временным откликом. Причиной «немгновенного» временного отклика может стать, например, накопление заряда в видиконе камеры, действующей с излишне высокой кадровой частотой. В изображающей системе без памяти наблюдаемое изображение может описываться выражением вида

$$F_O^{(i)}(x, y, t) = \mathcal{O}_P\{C, x, y, \lambda; C(x, y, t, \lambda)\}. \quad (13.1.2)$$

В другом особом случае временной отклик определяется взвешенным интегрированием по координате t на ограниченном временном интервале светового поля:

$$F_O^{(i)}(x, y, t) = \mathcal{O}_P\left\{C, x, y, \lambda; \int_{t-T}^{t} C(x, y, t, \lambda) L_i(t)\, dt\right\}, \quad (13.1.3)$$

где T — интервал интегрирования, $L_i(t)$ — временной импульсный отклик.

Во многих изображающих системах поэлементная спектральная характеристика представляет собой результат взвешенного интегрирования по λ входного светового поля. Таким образом, имеем

$$F_O^{(i)}(x, y, t) = \mathcal{O}_P\left\{C, x, y, t; \int_{0}^{\infty} C(x, y, t, \lambda) S_i(\lambda)\, d\lambda\right\}, \quad (13.1.4)$$

где $S_i(\lambda)$ — спектральная характеристика датчика.

Искажения часто моделируют двумя последовательными звеньями, описывающими пространственные эффекты и поэлементные эффекты, связанные с изменением интенсивности. Представим, например, что сначала возникают поэлементные искажения $\mathcal{O}_C\{\cdot\}$, а затем — пространственные искажения $\mathcal{O}_S\{\cdot\}$. В этом случае можно записать

$$F_O^{(i)}(x, y, t) = \mathcal{O}_S\{x, y, t, \lambda; \mathcal{O}_C\{C, t, \lambda; C(x, y, t, \lambda)\}\}. \quad (13.1.5)$$

При обратном порядке возникновения искажений запишем

$$F_O^{(i)}(x, y, t) = \mathcal{O}_C\{C, t, \lambda; \mathcal{O}_S\{x, y, t, \lambda; C(x, y, t, \lambda)\}\}. \quad (13.1.6)$$

В обоих случаях допускается зависимость пространственных и поэлементных искажений от длины волны и времени.

Линейные пространственные искажения можно представить как результат операции суперпозиции. На выходе изображающей системы без памяти с разделимыми поэлементными и пространственными искажениями будет наблюдаться изображение вида

$$F_O^{(i)}(x, y, t) = \mathcal{O}_C\left\{C, \lambda; \int_{-\infty}^{\infty}\int C(\alpha, \beta, t, \lambda) J(x, y, \alpha, \beta, t, \lambda)\, d\alpha\, d\beta\right\}, \tag{13.1.7}$$

где $J(\cdot)$ — импульсный отклик, описывающий пространственные искажения изображения. В общем случае импульсный отклик является функцией всех четырех пространственных координат x, y, α, β и длины волны. Такую систему называют линейной пространственно-зависимой. Особый случай представляют линейные пространственно-инвариантные искажения. В этом случае наблюдаемое изображение и входное световое поле связаны интегралом свертки:

$$F_O^{(i)}(x, y, t) = \mathcal{O}_C\left\{C, \lambda; \int_{-\infty}^{\infty}\int C(\alpha, \beta, t, \lambda) J(x-\alpha; y-\beta; t, \lambda)\, d\alpha\, d\beta\right\}. \tag{13.1.8}$$

В обобщенной модели на рис. 13.1.1 предполагается, что каждое из наблюдаемых изображений $F_O^{(i)}(x, y, t)$ переводится в цифровую форму методами, описанными в части 3; в результате в каждый момент времени t формируется массив отсчетов изображения $F_S^{(i)}(m_1, m_2, t)$. Отсчеты на выходе цифрового преобразователя связаны с входным наблюдаемым изображением соотношением

$$F_S^{(i)}(m_1, m_2, t) = \mathcal{O}_G\{F_O, x, y; F_O^{(i)}(x, y, t)\}, \tag{13.1.9}$$

где $\mathcal{O}_G\{\cdot\}$ — оператор, моделирующий процесс преобразования изображения в цифровую форму.

Система цифровой реставрации изображений вырабатывает выходной массив $F_K^{(i)}(k_1, k_2, t)$, осуществляя преобразование вида

$$F_K^{(i)}(k_1, k_2, t) = \mathcal{O}_K\{F_S^{(i)}(m_1, m_2, t)\}, \tag{13.1.10}$$

где $\mathcal{O}_K\{\cdot\}$ — оператор реставрации. Дисплей интерполирует выходные отсчеты системы цифровой реставрации и формирует непрерывное изображение $\hat{F}_I^{(i)}(x, y, t)$ — оценку исходного изображения. Для этой операции можно записать

$$\hat{F}_I^{(i)}(x, y, t) = \mathcal{O}_D\{F_S, x, y; F_K^{(i)}(k_1, k_2, t)\}, \tag{13.1.11}$$

где $\mathcal{O}_D\{\cdot\}$ — преобразование, выполняемое дисплеем.

Система цифровой реставрации изображений должна компенсировать искажения, вносимые изображающей системой, цифровым преобразователем и дисплеем, чтобы получить оценку гипотетического идеального изображения $F_I^{(i)}(x, y, t)$, воспроизводимого в системе, все элементы которой являются идеальными. Идеальная изображающая система способна сформировать идеальное изображение, описываемое функцией

$$F_I^{(i)}(x, y, t) = \int_0^{\infty} \int_{t-T}^{t} C(x, y, t, \lambda) U_i(\lambda) \, dt \, d\lambda, \qquad (13.1.12)$$

где $U_i(\lambda)$ — желательная спектральная характеристика, T — интервал наблюдения.

Обычно не удается осуществить идеальную реставрацию наблюдаемого изображения, после которой выходное изображение совпадает с исходным изображением. Цель проектирования системы реставрации изображений — минимизировать некоторую меру ошибки, определяемой различием $F_I^{(i)}(x, y, t)$ и $\hat{F}_I^{(i)}(x, y, t)$. В данной главе в основном будут рассматриваться методы, минимизирующие среднеквадратическую ошибку, т. е. разность функций исходного и реставрированного изображений:

$$\mathscr{E}_i = E\{[F_I^{(i)}(x, y, t) - \hat{F}_I^{(i)}(x, y, t)]^2\}. \qquad (13.1.13)$$

Часто желательно наложить дополнительные ограничения на процедуру минимизации ошибки; например, можно потребовать, чтобы функция, описывающая реставрированное изображение, имела только положительные значения, если изображение представляется распределением яркости.

Поскольку реставрация изображений должна осуществляться цифровым способом, часто более удобно находить меру ошибки только в дискретных точках идеального и оцененного изображений. Соответствующие дискретные массивы можно получить, пользуясь математическими моделями идеальных цифровых преобразователей:

$$F_D^{(i)}(n_1, n_2, t) = F_I^{(i)}(x, y, t)\,\delta(x - n_1 \Delta I, y - n_2 \Delta I) \qquad (13.1.14\text{а})$$

и

$$\hat{F}_D^{(i)}(n_1, n_2, t) = \hat{F}_I^{(i)}(x, y, t)\,\delta(x - n_1 \Delta I, y - n_2 \Delta I). \qquad (13.1.14\text{б})$$

Предполагается, что непрерывные изображения дискретизируются с шагом ΔI, удовлетворяющим критерию Найквиста. Кроме того, ошибку квантования считают пренебрежимо малой. Следует заметить, что процессы, представленные на рис. 13.1.1 блоками над пунктирной разделительной линией, отображают математические модели, а не физические операции, совершаемые над реальными

полями, представляющими изображения, и массивами их отсчетов. При указанной дискретизации непрерывных идеального и оцененного изображений соответствующая среднеквадратическая ошибка реставрации принимает вид

$$\mathscr{E}_i = E\{[F_D^{(i)}(n_1, n_2, t) - \hat{F}_D^{(i)}(n_1, n_2, t)]^2\}. \qquad (13.1.15)$$

Установление количественных соотношений, представленных на рис. 13.1.1, позволяет сформулировать проблему реставрации изображений в следующем виде:

При заданном дискретизованном наблюдаемом массиве изображения $F_S^{(i)}(m_1, m_2, t)$, полученном из исходного светового поля $C(x, y, t, \lambda)$, определить передаточную функцию $\mathscr{O}_K\{\cdot\}$, минимизирующую меру различия $F_I^{(i)}(n_1, n_2, t)$ и $\hat{F}_I^{(i)}(n_1, n_2, t)$ с учетом требуемых ограничений.

Исключительная сложность задачи реставрации изображений в формулировке, представленной выше, не позволила найти решение общего вида. Чтобы продолжить изложение, нам придется конкретизировать тип искажений и используемый метод реставрации. В последующих разделах описываются модели всех элементов обобщенной изображающей системы, представленной на рис. 13.1.1.

13.2. МОДЕЛИ ОПТИЧЕСКИХ СИСТЕМ

К числу крупных достижений оптики за последние 20 лет следует отнести применение системного подхода к исследованию оптических изображающих систем. Оптические приборы, состоящие из объективов, зеркал, призм и других подобных элементов, можно считать устройствами, осуществляющими детерминированное преобразование входного пространственного распределения света в некоторое выходное пространственное распределение света. Более того, системный подход можно обобщить, учитывая эффект распространения света в свободном пространстве или в некоторой диэлектрической среде.

В геометрической оптике предполагается, что в однородной среде световые лучи всегда распространяются прямолинейно. При таком допущении пучок лучей, попадающих на экран после прохождения через отверстие, формирует геометрическую проекцию этого отверстия. Однако при внимательном рассмотрении распределения света в области между светлой и темной зонами на экране оказывается, что граница не является резкой. Этот эффект тем сильнее, чем меньше размеры отверстия. Когда отверстие становится точечным, создается впечатление, что весь экран

освещен диффузным светом. Таким образом, возникает отклонение лучей света, называемое дифракцией. Количественные характеристики дифракции можно получить, если свет считать электромагнитным излучением, удовлетворяющим уравнениям Максвелла. Изложение полной теории оптических изображающих систем с позиций фундаментальных принципов электромагнитной теории дифракции — задача сложная и трудоемкая. После-

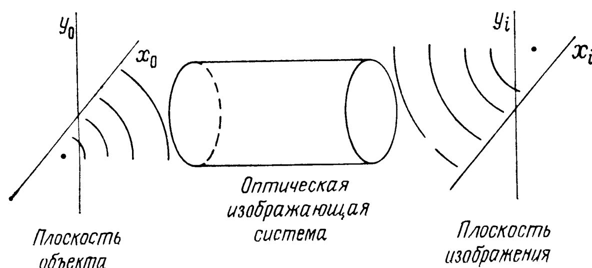


Рис. 13.2.1. Обобщенная оптическая изображающая система.

дующий материал содержит лишь ключевые моменты этой теории; подробные сведения можно получить в работах [1—3].

На рис. 13.2.1 представлена схема обобщенной оптической изображающей системы. Точка на плоскости объекта с координатами (x_o, y_o), имеющая интенсивность $I_o(x_o, y_o)$, излучает энергию в направлении изображающей системы, которая характеризуется входным и выходным зрачками и промежуточной средой, осуществляющей преобразование. Электромагнитные волны, испускаемые оптической системой, фокусируются в точке (x_i, y_i) плоскости изображения, в которой возникает интенсивность $I_i(x_i, y_i)$. Говорят, что изображающая система является дифракционно ограниченной, если распределение света на плоскости изображения, создаваемое точечным объектом, имеет вид сходящейся сферической волны, ограниченной только контуром выходного зрачка. Если волновой фронт электромагнитного излучения, испускаемого выходным зрачком, не является сферическим, говорят, что оптическая система имеет аберрацию.

В большинстве оптических систем формирования изображения световое излучение, испускаемое объектом, является результатом пропускания или отражения света некогерентного источника. Во многих случаях излучение, формирующее изображение, может считаться квазимонохроматическим в том смысле, что ширина его спектра мала по сравнению с центральной частотой излучения. В рамках этих двух допущений изображающая система, показанная на рис. 13.2.1, будет вести себя как линейная

система относительно интенсивности входного и выходного полей. В оптической системе зависимость между интенсивностями изображения и объекта может быть выражена интегралом суперпозиции вида

$$I_i(x_i,\ y_i)=\int\limits_{-\infty}^{\infty}\int H(x_i,\ y_i;\ x_o,\ y_o)\,I_o(x_o,\ y_o)\,dx_o\,dy_o, \qquad (13.2.1)$$

где $H(x_i,\ y_i;\ x_o,\ y_o)$ — импульсный отклик системы, т. е. распределение интенсивности в плоскости изображения, создаваемое точечным источником света. Импульсный отклик часто оказывается пространственно-инвариантным, и тогда зависимость «выход—вход» выражается интегралом свертки

$$I_i(x_i,\ y_i)=\int\limits_{-\infty}^{\infty}\int H(x_i-x_o,\ y_i-y_o)\,I_o(x_o,\ y_o)\,dx_o\,dy_o. \qquad (13.2.2)$$

В этом случае нормированные спектры Фурье

$$\mathcal{I}_o(\omega_x,\ \omega_y)=\frac{\int\limits_{-\infty}^{\infty}\int I_o(x_o,\ y_o)\exp\left[-i\,(\omega_x x_o+\omega_y y_o)\right]dx_o\,dy_o}{\int\limits_{-\infty}^{\infty}\int I_o(x_o,\ y_o)\,dx_o\,dy_o} \qquad (13.2.3\text{а})$$

и

$$\mathcal{I}_i(\omega_x,\ \omega_y)=\frac{\int\limits_{-\infty}^{\infty}\int I_i(x_i,\ y_i)\exp\left[-i\,(\omega_x x_i+\omega_y y_i)\right]dx_i\,dy_i}{\int\limits_{-\infty}^{\infty}\int I_i(x_i,\ y_i)\,dx_i\,dy_i} \qquad (13.2.3\text{б})$$

полей интенсивности объекта и изображения соответственно связаны следующим образом:

$$\mathcal{I}_o(\omega_x,\ \omega_y)=\mathcal{H}(\omega_x,\ \omega_y)\,\mathcal{I}_i(\omega_x,\ \omega_y), \qquad (13.2.4)$$

где $\mathcal{H}(\omega_x,\ \omega_y)$, часто называемая оптической передаточной функцией (ОПФ), определяется выражением вида

$$\mathcal{H}(\omega_x,\ \omega_y)=\frac{\int\limits_{-\infty}^{\infty}\int H(x,\ y)\exp\left[-i\,(\omega_x x+\omega_y y)\right]dx\,dy}{\int\limits_{-\infty}^{\infty}\int H(x,\ y)\,dx\,dy}. \qquad (13.2.5)$$

Модуль ОПФ $|\mathscr{H}(\omega_x, \omega_y)|$ называется частотно-контрастной характеристикой (ЧКХ) оптической системы.

Круглая тонкая линза является самой распространенной оптической системой формирования изображения. На рис. 13.2.2 показаны сечения ОПФ для такой линзы при расфокусировке различной степени [1, стр. 486, 4]. Большая расфокусировка приводит к тому, что в некоторой области пространственных частот ОПФ принимает отрицательные значения. В этом случае линза вызывает обращение контраста: темные объекты будут казаться светлыми, а светлые — темными.

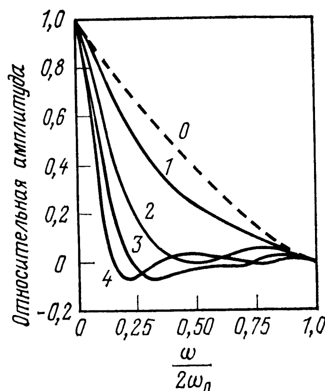


Рис. 13.2.2. Сечения оптической передаточной функции линзы при расфокусировке различной степени.

Земную атмосферу также можно рассматривать как оптическую систему. Показатель преломления нормальной атмосферы сохраняется приблизительно постоянным в пределах оптической протяженности объекта. В некоторых случаях, однако, возникает турбулентность атмосферы, при которой показатель преломления становится пространственно-зависимым, а результирующее изображение объекта — нерезким. Эквивалентный импульсный отклик имеет вид

$$H(x, y) = K_1 \exp[-(K_2x^2 + K_3y^2)^{5/6}], \tag{13.2.6}$$

где K_i — постоянные, которые были предсказаны математически, а затем проверены экспериментально [5] в случае формирования изображения с большой выдержкой. Чтобы упростить анализ, показатель степени $^5/_6$ часто заменяют единицей. При этом получают модель импульсного отклика гауссовой формы:

$$H(x, y) = K \exp\left\{-\left(\frac{x^2}{2b_x^2} + \frac{y^2}{2b_y^2}\right)\right\}, \tag{13.2.7}$$

где K — нормирующий коэффициент, b_x и b_y — коэффициенты пространственной нерезкости.

13.3. МОДЕЛИ ФОТОГРАФИЧЕСКИХ ПРОЦЕССОВ

Фотографическая запись изображений может осуществляться с использованием самых различных материалов и химических процессов. В данном разделе мы не претендуем ни на обзор фотографических методов, ни на углубленное рассмотрение физики фотографических процессов. Все это содержится в работах [6—8].

Однако здесь сделана попытка создания математических моделей фотографических процессов для описания фотографических элементов изображающей системы.

13.3.1. ОДНОЦВЕТНАЯ ФОТОГРАФИЯ

Для фотографической записи изображений чаще всего используют галогенидосеребряную эмульсию (рис. 13.3.1). В этом материале зерна галогенида серебра взвешены в прозрачном слое желатина на стеклянной, ацетатной или бумажной подложке. Негатив и диапозитив имеют прозрачную подложку; подложка отражающего фотоотпечатка изготовляется из белой бумаги. Когда на зерно попадает свет, происходит электрохимическая реакция и некоторая часть его превращается в металлическое серебро. Как говорят в этом случае, в зерне существует центр проявления. В процессе проявления химический проявитель действует на зерна с частичным содержанием серебра таким образом, что они целиком превращаются в частички металлического серебра. Фиксирование фотослоя заключается в удалении неэкспонированных зерен химическим способом.

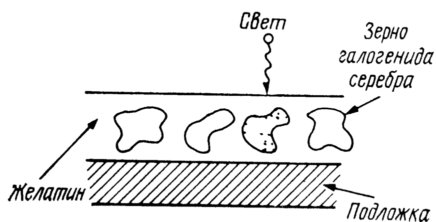


Рис. 13.3.1. Поперечное сечение фотопленки с галогенидосеребряной эмульсией.

Описанный фотографический процесс называется необратимым. С его помощью формируется изображение, являющееся негативным в том смысле, что содержание серебряных частичек обратно пропорционально интенсивности экспонирующего света. Получить позитивный фотоотпечаток позволяет двухступенчатый процесс с использованием необратимых материалов. Сначала изготовляют прозрачный негатив, а затем экспонируют через него фотобумагу. Результирующая плотность почернения проявленной фотобумаги пропорциональна интенсивности света, проходящего через негатив.

Обратимая фотопленка позволяет изготовить диапозитив. Такая фотопленка после экспонирования сначала обрабатывается фактически так же, как и необратимая фотопленка. При этом все зерна, на которые воздействовал свет, целиком превращаются в металлическое серебро. Затем все частички металлического серебра удаляют химическим способом. После этого фотопленка равномерно засвечивается, чтобы экспонировать оставшиеся зерна галогенида серебра (возможно также применение химического

процесса для этой цели). Вновь экспонированные зерна проявляют и фиксируют, в результате чего получается диапозитив с оптической плотностью, пропорциональной интенсивности света при исходном экспонировании.

Взаимосвязь между интенсивностью света, экспонирующего фотоматериал, и оптической плотностью негатива (диапозитива) или фотоотпечатка можно количественно описать с помощью сенситометрических величин. Ставится задача — разработать модель, которая позволяла бы предсказать спектральное распределение светового излучения, прошедшего через негатив (диапозитив) или отраженного от фотоотпечатка, при заданном спектральном распределении экспонирующего света и определенных физических параметрах фотографического процесса. Первый этап этого процесса — экспонирование зерен галогенида серебра — можно в первом приближении промоделировать интегральным выражением вида

$$X(C) = k_x \int C(\lambda) L(\lambda) \, d\lambda. \qquad (13.3.1)$$

Здесь $X(C)$ — интегральная экспозиция; $C(\lambda)$ — спектральная плотность энергии экспонирующего света; $L(\lambda)$ — спектральная чувствительность фотопленки или фотобумаги с учетом любых спектральных потерь, вносимых светофильтрами или оптическими элементами; k_x — экспозиция, регулируемая путем установки диафрагмы и выдержки. Выражению (13.3.1) соответствует условие фиксированной выдержки. В идеальном случае увеличение выдержки в какое-то число раз должно приводить к такому же увеличению экспозиции. К сожалению, такая зависимость точно не выполняется. Отклонение от линейной зависимости называют нарушением фотографического закона взаимозаменяемости. Другая аномалия в предсказании экспозиции — эффект перемежающихся экспозиций, суть которого состоит в том, что экспозиции для света с постоянной интенсивностью и прерывистого света с той же средней интенсивностью разные. Таким образом, пользоваться выражением (13.3.1) можно только с учетом его ограничений: оно является совершенно строгим в случае фиксированной экспозиции и освещения с постоянной интенсивностью.

Спектральный коэффициент пропускания $\tau(\lambda)$ проявленного негатива или диапозитива в идеальном случае связан с оптической плотностью экспоненциальным законом поглощения вида

$$\tau(\lambda) = \exp\{-d_e D(\lambda)\}, \qquad (13.3.2)$$

где $D(\lambda)$ — оптическая плотность в функции длины волны для опорной экспозиции, d_e — переменная, пропорциональная фактической экспозиции. В случае одноцветных фотоматериалов

кривая оптической плотности $D(\lambda)$ находится примерно на одном уровне во всей видимой части спектра. Из формулы (13.3.2) следует, что высоким оптическим плотностям соответствуют небольшие коэффициенты пропускания, и наоборот. На практике обычно выбирают в (13.3.2) такой коэффициент пропорциональности, чтобы измерения производились в единицах, выраженных сте-

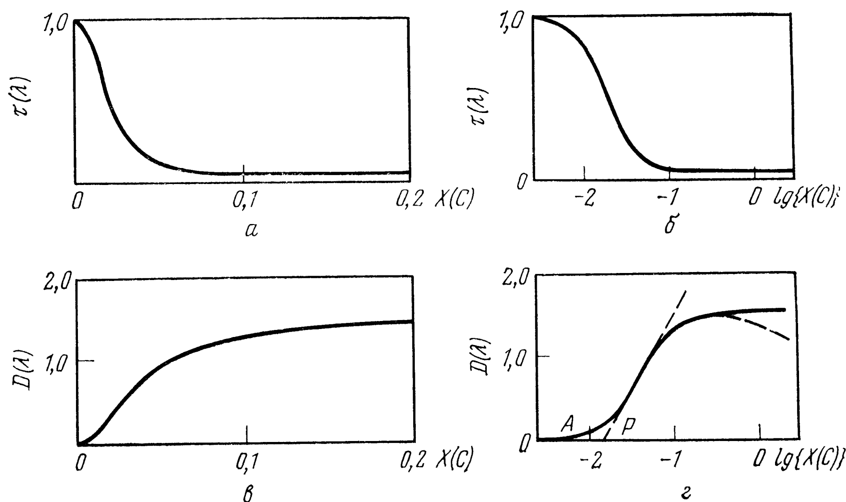


Рис. 13.3.2. Коэффициент пропускания $\tau(\lambda)$ и оптическая плотность $D(\lambda)$ в зависимости от экспозиции $X(C)$ для необратимой фотопленки.

пенью десяти. В этом случае выражение для коэффициента пропускания можно записать в виде

$$\tau(\lambda) = 10^{-\{d_x D(\lambda)\}}, \qquad (13\ 3.3)$$

где d_x — переменная, обратно пропорциональная экспозиции. Из формулы (13.3.3) следует, что оптическая плотность и коэффициент пропускания связаны логарифмическим законом. Таким образом, можно записать

$$d_x D(\lambda) = -\lg\{\tau(\lambda)\}. \qquad (13.3.4)$$

Спектральный коэффициент отражения $r_o(\lambda)$ фотоотпечатка также связан с плотностью почернения экспоненциальным за-

коном поглощения (13.3.2). Из соотношений (13.3.3) и (13.3.4) непосредственно получаем

$$r_o(\lambda) = 10^{-\{d_x D(\lambda)\}} \tag{13.3.5}$$

и

$$d_x D(\lambda) = \lg \{r_o(\lambda)\}, \tag{13.3.6}$$

где d_x — соответствующим образом вычисленная переменная, пропорциональная экспозиции фотобумаги.

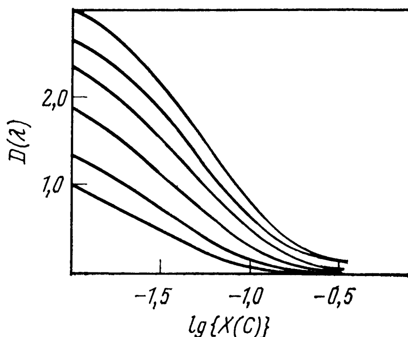


Рис. 13.3.3. Характеристические кривые для обратимой фотопленки в зависимости от времени проявления.

Модель отношений, связывающая оптическую плотность с коэффициентом пропускания или отражения, прозрачна и относительно точна. Следующий этап — моделирование зависимости интегральной экспозиции X (C) от переменной плотности d_x — сопряжен с большими трудностями. На рис. 13.3.2, *а* показана типичная кривая зависимости коэффициента пропускания негатива из необратимого фотоматериала от экспозиции. Следует отметить существенную нелинейность кривой во всем диапазоне экспозиций, за исключением относительно узкой области малых значений. На рис. 13.3.2,*б* эта же кривая представлена в логарифмическом масштабе по оси экспозиций. Как видно, между коэффициентом пропускания и логарифмом экспозиции существует примерно линейная зависимость, однако в указанной области значений экспозиции изображающие системы обычно практически не работают. В фотографии обычно интересуются переменной оптической плотности d_x, которая представлена кривыми на рис. 13.3.2, *в* и *г* в зависимости от экспозиции и логарифма экспозиции соответственно. Кривая зависимости оптической плотности от логарифма экспозиции называется характеристической кривой. На рис. 13.3.3 приведены характеристические кривые обратимой фотопленки. Приблизительно линейный участок кривой на рис. 13.3.2, *г* можно аппроксимировать прямой

$$d_x = \gamma [\lg \{X(C)\} - K_F], \tag{13.3.7}$$

где γ — коэффициент, характеризующий наклон прямой, K_F — значение логарифма экспозиции в точке пересечения оси абсцисс этой прямой. Угловой коэффициент прямой γ является мерой контраста фотопленки, а коэффициент K_F — мерой ее светочувствительности; другими словами, K_F характеризует минималь-

ную экспозицию, которая требуется для получения негатива с использованием линейной области характеристической кривой. Если диапазон экспозиций ограничить этой линейной областью, то можно, подставив выражение (13.3.7) в (13.3.3), получить

$$\tau(\lambda) = K_\tau(\lambda)\,[X(C)]^{-\gamma D(\lambda)}, \tag{13.3.8}$$

где

$$K_\tau(\lambda) \equiv 10^{\gamma K_F D(\lambda)}.$$

Пользуясь моделью экспозиции (13.3.1), моделями коэффициентов пропускания и отражения (13.3.3) и (13.3.5) и характеристической кривой [или ее линеаризированной моделью (13.3.7)], можно разработать математическую модель одноцветного фотографического процесса. В качестве примера такого моделирования рассмотрим процесс изготовления одноцветного диапозитива, который обеспечивал бы идеальное воспроизведение яркости светового излучения со спектральной плотностью $C(\lambda)$, которая определяется выражением вида

$$L(C) = \int C(\lambda)\, V(\lambda)\, d\lambda, \tag{13.3.9}$$

где $V(\lambda)$ — относительная спектральная световая эффективность При освещении диапозитива белым светом со спектральной плот ностью $E(\lambda)$ образуется световое излучение со спектральной плотностью $C_T(\lambda) = E(\lambda)\,\tau_T(\lambda)$ и яркостью

$$L(C_T) = \int \tau_T(\lambda)\, E(\lambda)\, V(\lambda)\, d\lambda. \tag{13.3.10}$$

Подставляя сюда выражение для коэффициента пропускания диапозитива (13.3.3), получаем

$$L(C_T) = \int 10^{-\{d_x D(\lambda)\}} E(\lambda)\, V(\lambda)\, d\lambda. \tag{13.3.11}$$

Параметр плотности d_x в этом выражении связан с интегральной экспозицией фотопленки $X(C)$ характеристической кривой. Если ширина интервала экспозиций и его положение таковы, что оказывается применимой линейная аппроксимация (13.3.7), то яркость диапозитива выразится как

$$L(C_T) = \int K_\tau(\lambda)\,[X(C)]^{-\gamma D(\lambda)}\, D(\lambda)\, V(\lambda)\, d\lambda. \tag{13.3.12}$$

В случае одноцветных диапозитивов характеристическая кривая обычно имеет примерно постоянный уровень во всем видимом спектре. Поэтому, полагая $D(\lambda) = 1$ и $K_\tau(\lambda) = K_\tau$, можно написать

$$L(C_T) = K_\tau\,[X(C)]^{-\gamma} L(E), \tag{13.3.13}$$

где $L(E)$ — яркость светового излучения, освещающего диапозитив. Путем выбора и соответствующей обработки фотопленки обеспечим коэффициент $\gamma = -1$ при отрицательном наклоне характеристической кривой. В этом случае яркость диапозитива будет прямо пропорциональна экспозиции. С учетом этого условия из формулы (13.3.1) получаем

$$L(C_T) = K_\tau L(E)\, k_x \int C(\lambda)\, L(\lambda)\, d\lambda, \qquad (13.3.14)$$

где $L(\lambda)$ — спектральная чувствительность фотоэмульсии. Сравнение формул (13.3.9) и (13.3.14) показывает, что яркость экспонирующего света и яркость диапозитива прямо пропорциональны, если характеристика спектральной чувствительности фотоэмульсии идентична кривой относительной спектральной световой эффективности источника. На практике первая кривая существенно отличается от второй, поэтому даже в идеализированных условиях фотографического процесса, принятых в модели, не удается обеспечить точное воспроизведение яркости. Дополнительные искажения, вносимые фотографическим процессом, обусловлены отклонением коэффициента γ фотопленки от единицы, использованием экспозиций вне линейной области характеристической кривой и другими факторами, которые даже не рассматривались при разработке описанной простой модели.

13.3.2. ЦВЕТНАЯ ФОТОГРАФИЯ

В современной цветной фотографии для изготовления диапозитивов или негативов используется трехслойная цветная пленка, поперечное сечение которой показано на рис. 13.3.4. Первый слой представляет собой галогенидосеребряную эмульсию, чувствительную к синему свету. Желтый светофильтр под синим слоем не позволяет синему излучению распространяться в направлении зеленого и красного эмульсионных слоев, также чувствительных к этому излучению. Эмульсионные слои находятся на прозрачной подложке. В процессе проявления синий эмульсионный слой превращается в слой желтого красителя, причем концентрация этого красителя пропорциональна экспозиции в синих лучах для негатива и обратно пропорциональна для диапозитива. Аналогично зеленый и красный эмульсионные слои превращаются в слои пурпурного и голубого красителей соответственно. Цветные отпечатки можно изготовить с использованием различных процессов [7]. Самый распространенный метод изготовления цветного фотоснимка основан на экспонировании необратимой цветной фотобумаги через цветной негатив.

При разработке математической модели цветного фотографического процесса можно считать, что каждый эмульсионный слой реагирует на световое излучение так же, как и эмульсионный слой одноцветного фотоматериала. Как первое приближение такое предположение является допустимым. Однако часто возникает сильное взаимодействие между слоями. Типичные характеристики спектральной чувствительности трех эмульсионных

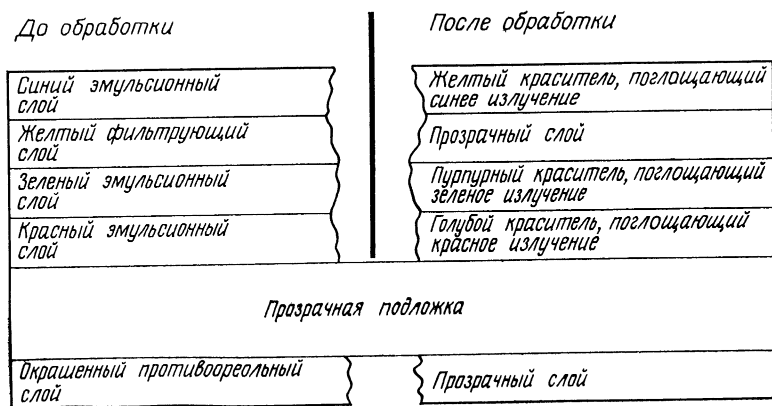


Рис. 13.3.4. Трехслойная цветная фотопленка.

слоев показаны на рис. 13.3.5. Интегральные экспозиции для слоев определяются выражениями вида

$$X_R(C) = d_R \int C(\lambda) L_R(\lambda)\, d\lambda, \qquad (13.3.15а)$$

$$X_G(C) = d_G \int C(\lambda) L_G(\lambda)\, d\lambda, \qquad (13.3.15б)$$

$$X_B(C) = d_B \int C(\lambda) L_B(\lambda)\, d\lambda, \qquad (13.3.15в)$$

где d_R, d_G, d_B — постоянные. Эти коэффициенты пропорциональности выбираются таким образом, чтобы при освещении опорным белым светом все три экспозиции оказывались равными и фотопленка не использовалась в режиме насыщения. В процессе химического проявления фотопленки формируется негатив с тремя поглощающими слоями красителей — голубого, пурпурного и желтого.

Коэффициент пропускания $\tau_T(\lambda)$ обработанного негатива является произведением коэффициентов пропускания $\tau_{TC}(\lambda)$,

$\tau_{TM}(\lambda)$, $\tau_{TY}(\lambda)$ голубого, пурпурного и желтого красителей соответственно. Следовательно,

$$\tau_T(\lambda) = \tau_{TC}(\lambda)\,\tau_{TM}(\lambda)\,\tau_{TY}(\lambda). \qquad (13.3.16)$$

Коэффициент пропускания каждого слоя красителя зависит от его спектральной характеристики поглощения и концентрации.

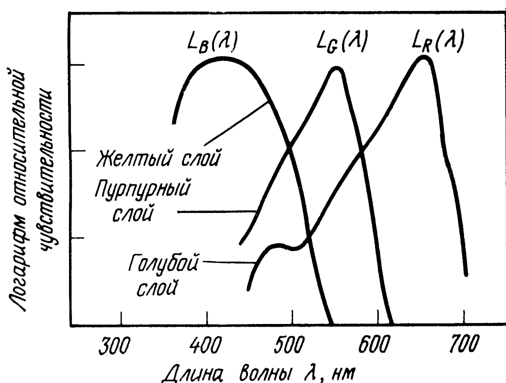


Рис. 13.3.5. Характеристики спектральной чувствительности типовых эмульсионных слоев цветной фотопленки.

Эту функциональную зависимость удобно выразить через относительную плотность. В этом случае для каждого из слоев получаем

$$\tau_{TC}(\lambda) = 10^{\{-cD_{NC}(\lambda)\}}, \qquad (13.3.17а)$$

$$\tau_{TM}(\lambda) = 10^{\{-mD_{NM}(\lambda)\}}, \qquad (13.3.17б)$$

$$\tau_{TY}(\lambda) = 10^{\{-yD_{NY}(\lambda)\}}, \qquad (13.3.17в)$$

где c, m, y — относительное содержание голубого, пурпурного и желтого красителей, $D_{NC}(\lambda)$, $D_{NM}(\lambda)$, $D_{NY}(\lambda)$ — спектральные оптические плотности для единичных количеств этих красителей. Коэффициент пропускания, рассчитанный для единичных количеств красителей, составит

$$\tau_{TN}(\lambda) = 10^{\{-D_{TN}(\lambda)\}}, \qquad (13.3.18)$$

где

$$D_{TN}(\lambda) = D_{NC}(\lambda) + D_{NM}(\lambda) + D_{NY}(\lambda).$$

Негатив или диапозитив с такими параметрами, освещенный опорным белым светом, будет выглядеть нейтрально серым. Типичные кривые спектральной оптической плотности трех красителей обратимой пленки представлены на рис. 13.3.6; здесь также показана соответствующая им кривая спектральной нейтральной оптической плотности.

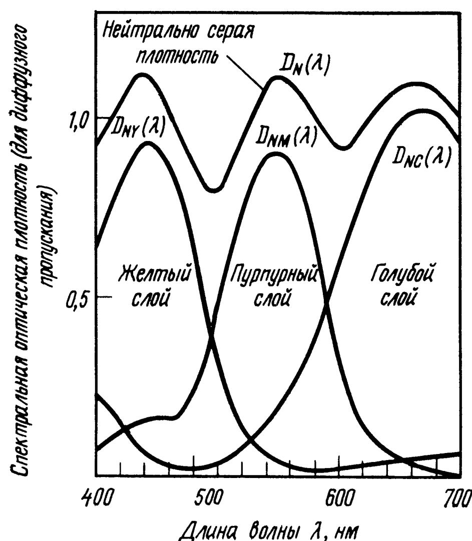


Рис. 13.3.6. Характеристики спектральной оптической плотности красителей и нейтрально серой плотности для типовой обратимой цветной фотопленки.

Между значениями экспозиции и плотностью слоев красителей в общем случае существует сложная зависимость. Например, количество голубого красителя, образующегося в случае нелинейной зависимости, определяется не только экспозицией в красных лучах, но и экспозициями в зеленых и голубых лучах, хотя и в меньшей степени. Аналогичные зависимости существуют между образующимися количествами пурпурного и желтого красителей и соответствующими экспозициями. Во многих случаях перекрестными связями между слоями (перекрестными изображениями) можно пренебречь и считать, что голубой краситель образуется только благодаря экспозиции в красном, пурпурный — в зеленом, синий — в желтом. При таком допущении соотношение «плотность—экспозиция» красителя может представляться характеристической кривой зависимости эквивалентной нейтрально серой плотности от логарифма экспозиции каждого красителя. На рис. 13.3.7 представлены типичные характеристические кривые для обратимой пленки. Центральные участки каждой из этих кривых можно аппроксимировать выражениями вида

$$c = K_{FC} + \gamma_C \lg(X_R), \tag{13.3.19а}$$

$$m = K_{FM} + \gamma_M \lg(X_G), \tag{13.3.19б}$$

$$y = K_{FY} + \gamma_Y \lg(X_B), \tag{13.3.19в}$$

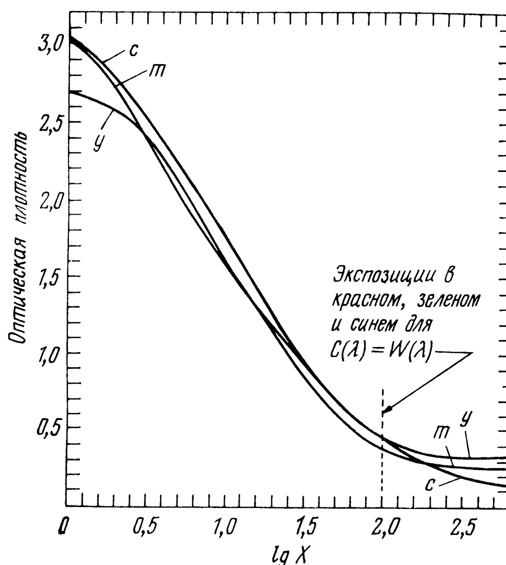


Рис. 13.3.7. Характеристические кривые для типовой обратимой цветной фотопленки.

Линейная аппроксимация $c = 3{,}45 - 1{,}66 \lg X_R$, $m = 3{,}10 - 1{,}51 \lg X_G$, $y = 3{,}25 - 1{,}62 \lg X_B$.

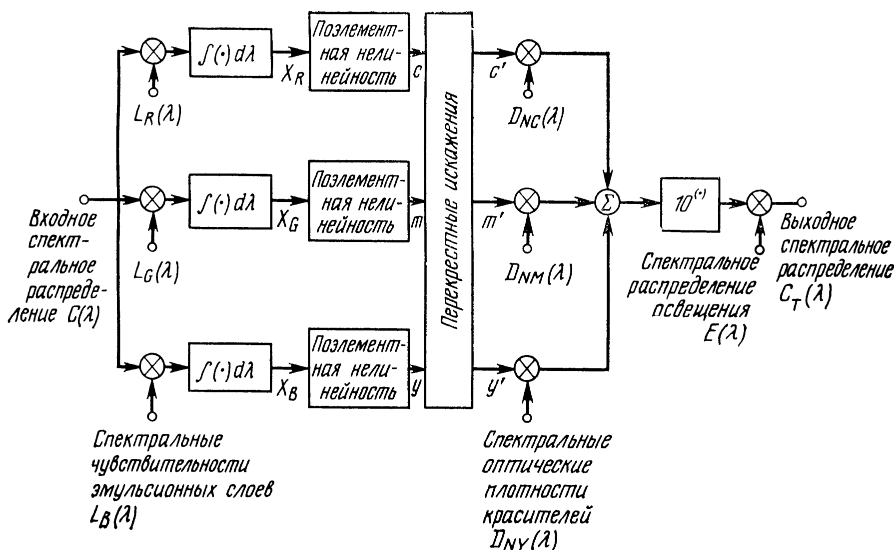


Рис. 13.3.8. Модель цветного фотографического процесса.

где γ_C, γ_M, γ_Y — тангенсы угла наклона показанных кривых в линейной области, называемые показателями степени γ слоев красителей.

Спектральная плотность светового излучения после прохождения через проявленный негатив (диапозитив) определяется произведением спектрального коэффициента пропускания на спектральную плотность падающего светового потока $E(\lambda)$:

$$C_T(\lambda) = E(\lambda)\, 10^{-\{cD_{NC}(\lambda) + mD_{NM}(\lambda) + yD_{NY}(\lambda)\}}, \qquad (13.3.20)$$

На рис. 13.3.8 показана схема модели полного цветного фотографического процесса записи—воспроизведения. Исходная спектральная плотность излучения $C(\lambda)$ и спектральная плотность излучения на выходе негатива (позитива) $C_T(\lambda)$, отнесенная к любому данному элементу, редко оказываются одинаковыми. Иначе говоря, в фотографическом процессе правильная передача спектрального состава излучения, как правило, не достигается. Более того, обычно не обеспечивается даже верное воспроизведение цветов, поскольку в общем случае координаты цвета излучения $[C_T]$

$$U(C_T) = \int C_T(\lambda)\, U_S(\lambda)\, d\lambda, \qquad (13.3.21а)$$

$$V(C_T) = \int C_T(\lambda)\, V_S(\lambda)\, d\lambda, \qquad (13.3.21б)$$

$$W(C_T) = \int C_T(\lambda)\, W_S(\lambda)\, d\lambda \qquad (13.3.21в)$$

не равны координатам цвета излучения $[C]$

$$U(C) = \int C(\lambda)\, U_S(\lambda)\, d\lambda, \qquad (13.3.22а)$$

$$V(C) = \int C(\lambda)\, V_S(\lambda)\, d\lambda, \qquad (13.3.22б)$$

$$W(C) = \int C(\lambda)\, W_S(\lambda)\, d\lambda. \qquad (13.3.23в)$$

В формулах (13.3.21) и (13.3.22) $U_S(\lambda)$, $V_S(\lambda)$, $W_S(\lambda)$ представляют собой координаты цвета спектральных цветов в равноконтрастной колориметрической системе. Реставрация цветных изображений рассмотрена в гл. 16.

13.4. МОДЕЛИ ЦИФРОВЫХ ПРЕОБРАЗОВАТЕЛЕЙ ИЗОБРАЖЕНИЯ

Как было указано в гл. 6, физический цифровой преобразователь может быть источником искажения в цифровой изображающей системе в том смысле, что непрерывное изображение, восстановленное по физическим отсчетам, может отличаться от непрерывного изображения, подвергнутого дискретизации. Боль-

шие искажения могут возникнуть вследствие образования ложных частот при дискретизации с недостаточно высокой частотой. Процедура исправления таких искажений описана в гл. 15. Здесь будет предполагаться, что частота дискретизации удовлетворяет критерию Найквиста (теореме отсчетов). Другой источник искажений — усечение интерполирующих функций на краях изображения, обусловленное ограниченностью массива отсчетов. Ошибки усечения обычно значительны только в краевой области изображения, ширина которой составляет 5—10 элементов.

Дискретизация импульсами конечной длительности (протяженности) приводит к эквивалентной нерезкости дискретизованного изображения. Как следует из формулы (4.2.10), массив отсчетов $F_S(m_1, m_2, t)$ и функция наблюдаемого изображения связаны интегралом свертки вида

$$F_S(m_1, m_2, t) = \int_{-\infty}^{\infty}\int F_O(\alpha, \beta, t)\, P(\alpha - m_1 \Delta x, \beta - m_2 \Delta y)\, d\alpha\, d\beta, \tag{13.4.1}$$

где $P(x, y)$ описывает форму дискретизирующего импульса, а Δx и Δy — шаги дискретизации. Эффект дискретизации импульсами конечной протяженности можно промоделировать операцией свертки непрерывного поля изображения, подлежащего дискретизации, с импульсным откликом $P(-x, -y)$, выполненной перед идеальной дискретизацией дельта-функциями Дирака.

Ошибки квантования являются источником сильных искажений изображения при использовании квантователя с недостаточным числом уровней квантования. Методы реставрации, предназначенные для компенсации ошибок квантования при ИКМ и ДИКМ, а также различные методы кодирования изображений с преобразованием представлены в части 6.

13.5. МОДЕЛЬ ДИСПЛЕЯ

Дисплей на рис. 13.1.1 воспроизводит непрерывное изображение, описываемое функцией $\hat{F}_I(x, y, t)$, интерполируя массив отсчетов $F_K(k_1, k_2, t)$, сформированный процессором цифровой реставрации, с помощью интерполирующей функции $R(x, y)$. Дисплеи часто обладают поэлементной нелинейностью, зависящей от величины F_K. В некоторых системах поэлементная нелинейность также является функцией координат (x, y) на плоскости воспроизведения. С учетом всех этих факторов выражение (13.1.11) дает обобщенную модель дисплея:

$$\hat{F}_I(x, y, t) = \mathcal{O}_D\{F_S, x, y; F_K(m_1, m_2, t)\}, \tag{13.5.1}$$

где $\mathcal{O}_D\{\cdot\}$ моделирует все виды преобразований, осуществляемых дисплеем.

Часто поэлементную нелинейность можно учитывать отдельно от операции интерполяции, вводя функцию

$$F_G(k_1, k_2, t) = \mathcal{O}_P\{F_K, x, y; F_K(k_1, k_2, t)\}, \qquad (13.5.2)$$

где $\mathcal{O}_P\{\cdot\}$ — оператор поэлементной нелинейности, которая затем используется при пространственной интерполяции, так что

$$\hat{F}_I(x, y, t) = \sum_{k_1=-K_1}^{K_1} \sum_{k_2=-K_2}^{K_2} F_G(k_1, k_2, t) R(x - k_1 \Delta x, y - k_2 \Delta y). \qquad (13.5.3)$$

Воспроизводимое изображение занимает область размера $(2K_1 + 1)\Delta x \times (2K_2 + 1)\Delta y$.

Идеальная функция пространственной интерполяции $R(x, y)$ должна иметь вид двумерной sinc-функции или двумерной бесселевой функции (см. разд. 4.1). К сожалению, такие интерполирующие функции имеют области отрицательных значений и, следовательно, не могут быть точно реализованы суммированием положительных световых интенсивностей. Результат использования неоптимальных интерполирующих функций можно оценить поэлементным сравнением функций воспроизводимого изображения $\hat{F}_I(x, y, t)$ и идеализированной модели воспроизводимого изображения $F_I(x, y, t)$, для чего эти изображения следует подвергнуть идеальной математической дискретизации дельта-функциями $\delta(x - n_1 \Delta I, y - n_2 \Delta I)$, образующими квадратную решетку с шагом ΔI. Введем для таких идеализированных массивов отсчетов обозначения $\hat{F}_D(n_1, n_2, t)$ и $\tilde{F}_D(n_1, n_2, t)$ соответственно. Связь между функциями $\hat{F}_D(n_1, n_2, t)$ и $F_G(k_1, k_2, t)$ можно представить как преобразование векторов:

$$\hat{\mathbf{f}}_D = \mathbf{R}\mathbf{f}_G, \qquad (13.5.4)$$

где $\hat{\mathbf{f}}_D$ — вектор размера $N^2 \times 1$, $\mathbf{f}_G$ — вектор размера $K^2 \times 1$, $\mathbf{R}$ — матрица отсчетов интерполирующей функции дисплея размера $N^2 \times K^2$. Аналогично вектор, представляющий изображение, сформированное при использовании оптимальной интерполирующей функции, есть

$$\tilde{\mathbf{f}}_D = \mathbf{R}_O\mathbf{f}_G, \qquad (13.5.5)$$

где $\mathbf{R}_O$ — матрица отсчетов оптимальной интерполирующей функции дисплея размера $N^2 \times K^2$.

В гл. 15 рассматриваются методы видоизменения элементов $\mathbf{f}_G$ в (13.5.5) с целью повышения верности воспроизведения.

13.6. МОДЕЛИ ШУМОВ ИЗОБРАЖЕНИЯ

Любое изображение искажается случайным образом из-за шумов датчика, зернистости фотопленки или флуктуаций оптических свойств атмосферы. Для удобства будем называть все такие эффекты шумами.

13.6.1. ШУМ ФОТОДЕТЕКТОРА

На рис. 13.6.1 представлена простая модель фотодетектора— датчика видеосигнала. Идеальный фотодетектор должен действовать как преобразователь интенсивности излучения в ток или, что эквивалентно, как преобразователь падающих фотонов в электроны. Генерируемый фотодетектором ток проходит через электрический фильтр и создает на нагрузочном резисторе выходной сигнал.

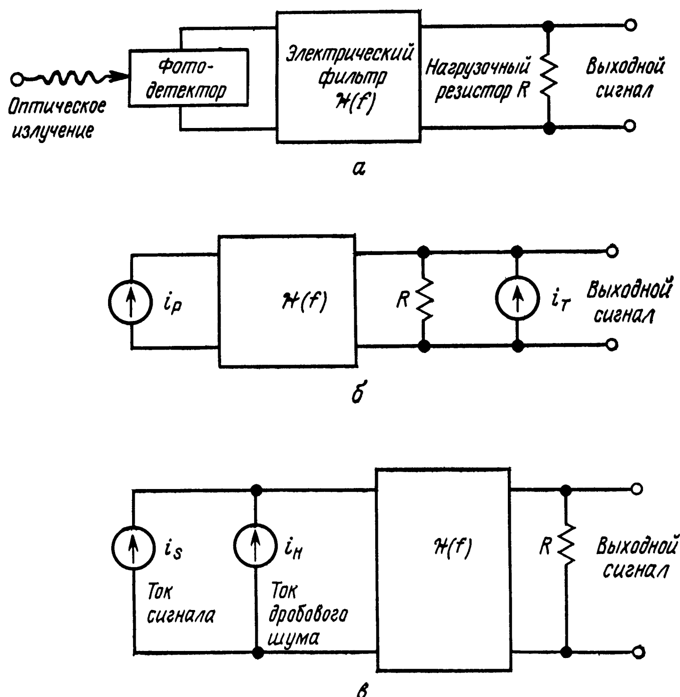


Рис. 13.6.1. Модели шумов фотодетектора.

а — схема фотодетектора; *б* — модель при наличии теплового шума; *в* — модель при наличии дробового шума.

Тепловой шум, вероятно, является наиболее общим видом шума, возникающего в электронных изображающих системах. Он является результатом случайных флуктуаций электронов в резистивных элементах фотодетекторов или усилителей, подключаемых к датчикам [9, 10]. Моделью теплового шума может служить аддитивный гауссовый случайный процесс, независимый от детектируемого изображения. Чтобы проиллюстрировать расчет теплового шума, рассмотрим модель фотодетектора на рис. 13.6.1, предполагая, что единственным существенным резистивным элементом схемы является нагрузочный резистор. На рис. 13.6.1,*б* фотодетектор заменен идеальным генератором тока i_P; ток, вырабатываемый этим генератором, прямо пропорционален интенсивности падающего света в некоторой точке изображения. Составляющую теплового шума можно промоделировать генератором тока i_T, включенным параллельно резистору. Этот ток представляет гауссовый случайный процесс с нулевым средним и дисперсией $\sigma_{i_T}^2 = N_T/R$, где N_T — мощность теплового шума на выходе системы. Легко показать [11, стр. 147], что

$$N_T = 2kT \int_{-\infty}^{\infty} |\mathscr{H}_E(f)|^2\, df, \tag{13.6.1}$$

где k — постоянная Больцмана, равная $1{,}38 \cdot 10^{-23}$ Дж/К, T — абсолютная температура, К, $\mathscr{H}_F(f)$ — эквивалентная частотная характеристика совокупности электрического фильтра и резистора нагрузки. В простейшем случае, когда фильтр представляет собой параллельное соединение конденсатора емкости C и нагрузочного резистора, мощность теплового шума равна $N_T = kT/RC$. При постоянном токе фотодетектора и линейном эквивалентном фильтре для плотности вероятности тока на выходе схемы можно непосредственно записать

$$p(i) = [2\pi\sigma_{i_T}^2]^{-1/2} \exp\left\{-\frac{[i - i_S]^2}{2\sigma_{i_T}^2}\right\}, \tag{13.6.2}$$

где i_S — эквивалентный средний ток сигнала на выходе схемы, соответствующий току фотодетектора i_P, протекающему через фильтр.

Даже при постоянной интенсивности падающего света ток фотодетектора не является строго постоянной величиной. Фотоэлектрическим датчикам свойственна неопределенность измерения, обусловленная квантовомеханической природой света. При малых уровнях освещенности число электронов, испускаемых фотодетектором, определяется плотностью распределения вероятностей Пуассона [12]. Для времени наблюдения τ распре-

деление вероятностей числа импульсов тока фотодетектора определяется выражением

$$\Pr\left(i_P = \frac{jq}{\tau}\right) = \frac{(u_S + u_H)^j \exp\{-(u_S + u_H)\}}{j!}, \qquad (13.6.3)$$

где j — положительное целое число, q — заряд электрона (равный $1{,}6\cdot 10^{-19}$ Кл), u_S — среднее число электронов, испускаемых фотодетектором при его освещении, u_H — среднее число испускаемых электронов, обусловленных темновым током фотодектора и посторонним фоновым излучением. Результирующая неопределенность в знании величины тока фотодетектора — это и есть так называемый дробовой шум в случае фотоэмиссионного детектора и шум генерации — рекомбинации в фотопроводящих и фотогальванических детекторах [13]. В фотодетекторе с достаточным внутренним усилением, таком, как фотоэлектронный умножитель, дробовой шум обычно преобладает над тепловыми шумами любого происхождения, возникающими в последующих цепях; в фотодетекторах другого типа, как правило, преобладает тепловой шум и фототок можно считать постоянной величиной. В любом случае, однако, при большом числе испускаемых фотоэлектронов распределение Пуассона (13.6.3) можно аппроксимировать гауссовым распределением [11, стр. 162] вида

$$p(i_p) = \left[\frac{2\pi q^2 u_S}{\tau^2}\right]^{-1/2} \exp\left\{\frac{-(i_p - qu_S/\tau)^2}{2q^2 u_S/\tau^2}\right\}. \qquad (13.6.4)$$

Таким образом, шум, вносимый большинством датчиков видеосигнала, можно моделировать случайными процессами с гауссовым или пуассоновым распределением. Шум с распределением Пуассона возникает только в том случае, когда изображение создается световым излучением чрезвычайно малой интенсивности, а фотодетектор обладает большим внутренним фотоэлектронным усилением. Другие практические фотодетекторные схемы обычно хорошо моделируются аддитивным гауссовым шумом, преобладающим над остальными шумами. Точный расчет дисперсии шума часто связан с анализом сложных электрических цепей, поэтому обычно прибегают к косвенным измерительным методам.

13.6.2. ШУМ ЗЕРНИСТОСТИ ФОТОПЛЕНКИ

В результате экспонирования и обработки фотопленки зерна галогенида серебра, подвергнутые достаточному световому воздействию, превращаются в зерна металлического серебра. Этот процесс не является чисто детерминированным: одинаковые галогенидосеребряные зерна при равной экспозиции не обязательно становятся серебряными зернами одинакового размера и одина

ковой формы; следует также учесть, что образованные зерна случайным образом распределены по площади обработанной пленки. Эта характерная для процесса формирования зерен серебра хаотичность, называемая шумом зернистости фотопленки, является причиной случайности или неопределенности количества света, проходящего через диапозитив или отражаемого фотоотпечатком [8, 14].

В классической модели [15, 16] флуктуационного шума зернистости предполагается гауссово распределение

$$p[D(x, y)] = [2\pi\sigma_D^2(x, y)]^{-1/2} \exp\left\{-\frac{[D(x, y) - u_D(x, y)]^2}{2\sigma_D^2(x, y)}\right\} \quad (13.6.5)$$

оптической плотности экспонированной фотопленки $D(x, y)$ со средним значением плотности $u_D(x, y)$ и дисперсией $\sigma_D^2(x, y)$. Средняя плотность измеряется путем пространственного усреднения в окрестности точки (x, y) в пределах некоторого окна, охватывающего множество отдельных зерен, а среднеквадратическое отклонение моделируется [17] выражением вида

$$\sigma_D(x, y) = \alpha\, [u_D(x, y)]^{\beta}, \quad (13.6.6)$$

где β — постоянная и

$$\alpha = 0{,}66 \left[\frac{a}{A}\right]^{1/2} \quad (13.6.7)$$

— эмпирически выбираемый параметр, причем a — средняя площадь зерна фотопленки, A — площадь обследуемой области фотопленки. Фалконер [17] и Хуанг [18] предложили для β значения $^1/_2$ и $^1/_3$ соответственно. Экспериментальные исследования показали, что пространственная корреляция флуктуаций содержания зерен весьма мала и, следовательно, шум зернистости фотопленки можно моделировать белым шумом. Плотность $D(x, y)$ с гауссовым распределеним (13.6.5) можно представить в виде суммы средней плотности $u_D(x, y)$ и гауссового случайного процесса с нулевым средним $N(x, y)$:

$$D(x, y) = u_D(x, y) + \sigma_D(x, y)\, N(x, y) \quad (13.6.8а)$$

или

$$D(x, y) = u_D(x, y) + \alpha\, [u_D(x, y)]^{\beta} N(x, y). \quad (13.6.8б)$$

Из этого соотношения видно, что шум зернистости, зависящий от записываемой информации, является аддитивным при его выражении через плотность [19]. На основе формулы (13.3.3) можно записать коэффициент пропускания экспонированной фотопленки при фиксированной длине волны как

$$\tau(x, y) = 10^{-\{D(x, y)\, D(\lambda)\}}, \quad (13.6.9)$$

где $D(\lambda)$ — спектральная оптическая плотность фотопленки. При этом спектральный коэффициент пропускания

$$\tau(x, y) = \tau_o(x, y)\, 10^{-\{\sigma_D(x, y)\, N(x, y)\, D(\lambda)\}} \qquad (13.6.10)$$

можно представить в виде произведения среднего коэффициента пропускания без учета зернистости $\tau_o(x, y)$ и множителя, пропор-

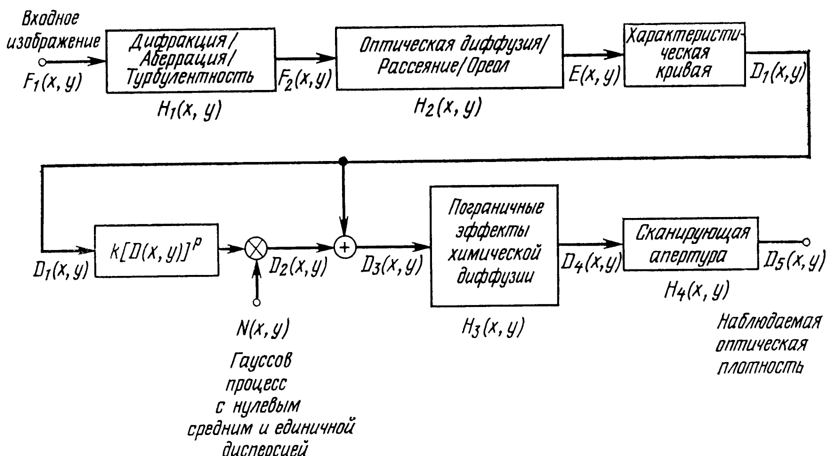


Рис. 13.6.2. Модель фотографической записи изображения [20].

ционального $\sigma_D(x, y)$. Таким образом, шум зернистости фотопленки, выраженный через интенсивность, моделируется мультипликативным шумовым процессом [19].

Надери и Савчук [20] разработали модель фотографической записи изображения, основанную на модели Келли [21]. В их модели (рис. 13.6.2) учтены пространственная нерезкость изображения, вносимая изображающей системой, фотографическим процессом и сканером, а также нелинейность фотопленки. Применение этой модели при реставрации изображений описано в гл. 15.

13.7. МОДЕЛИ ЦИФРОВОЙ РЕСТАВРАЦИИ ИЗОБРАЖЕНИЙ

В начале этой главы были введены обобщенная модель изображающей системы и модель процессора цифровой реставрации изображений. Затем на базе обобщенной модели были описаны и промоделированы типичные элементы изображающей системы. Теперь рассмотрим несколько моделей цифровой реставрации изображений. Эти модели будут использованы в последующих главах для описания ряда методов реставрации изображений.

Процесс реставрации можно описать с помощью следующих выражений:

$$F_S(m_1, m_2) = \mathcal{O}_M\{F_I(n_1, n_2), N_1(m_1, m_2), \ldots, N_N(m_1, m_2)\}, \quad (13.7.1а)$$

$$F_K(k_1, k_2) = \mathcal{O}_R\{F_S(m_1, m_2)\}, \quad (13.7.1б)$$

$$\hat{F}_I(n_1, n_2) = \mathcal{O}_D\{F_K(k_1, k_2)\}. \quad (13.7.1в)$$

Здесь F_S — массив наблюдаемых отсчетов изображения, F_I — массив элементов идеально дискретизованного изображения, $\hat{F}_I$ — массив элементов идеально дискретизованной оценки изображения, F_K — массив компенсированных отсчетов на выходе системы цифровой реставрации, N_i — массивы отсчетов составляющих шума, возникающих в различных элементах системы, $\mathcal{O}_M(\cdot)$, $\mathcal{O}_R(\cdot)$, $\mathcal{O}_D(\cdot)$ — обобщенные передаточные функции изображающей системы, реставрирующего процессора и дисплея соответственно.

Соотношения (13.7.1) можно представить в эквивалентной векторной форме:

$$\mathbf{f}_S = \mathcal{O}_M\{\mathbf{f}_I, \mathbf{n}_1, \ldots, \mathbf{n}_N\}, \quad (13.7.2а)$$

$$\mathbf{f}_K = \mathcal{O}_R\{\mathbf{f}_S\}, \quad (13.7.2б)$$

$$\hat{\mathbf{f}}_I = \mathcal{O}_D\{\mathbf{f}_K\}, \quad (13.7.2в)$$

где $\mathbf{f}_S$ — векторы, полученные разверткой по столбцам соответствующих массивов. В последующих главах рассмотрены некоторые методы нахождения приближенных решений (13.7.1) или (13.7.2). К сожалению, решения общего вида до сих пор не получены; поэтому мы обратимся к частным решениям.

Модель, показанная на рис. 13.7.1, является наиболее общей моделью цифровой реставрации изображений, в которой изображение приобретает пространственные искажения, приводящие к ухудшению резкости, электрический датчик реагирует нелинейным образом на интенсивность входного излучения, а усилитель датчика вносит гауссов шум, не зависящий от изображения. Затем следует цифровой преобразователь, который также может увеличить нерезкость дискретизованного изображения в случае использования дискретизирующих импульсов конечной протяженности.

Возникновение нерезкости изображения можно промоделировать операцией суперпозиции с импульсным откликом $J(x, y)$, который может быть пространственно-зависимым. Предполагается, что нелинейная реакция датчика на изображение $F_B(x, y)$ имеет поэлементный характер и что на выходе датчика имеется

аддитивное шумовое поле $N(x, y)$. Эффект дискретизации импульсами конечной протяженности симметричной формы $P(x, y)$ моделируется операцией свертки $F_O(x, y)$ с $P(x, y)$ и последующей идеальной дискретизацией.

Цель реставрации состоит в формировании массива отсчетов $\hat{F}_I(r_1, n_2)$, являющихся оценками отсчетов идеального входного

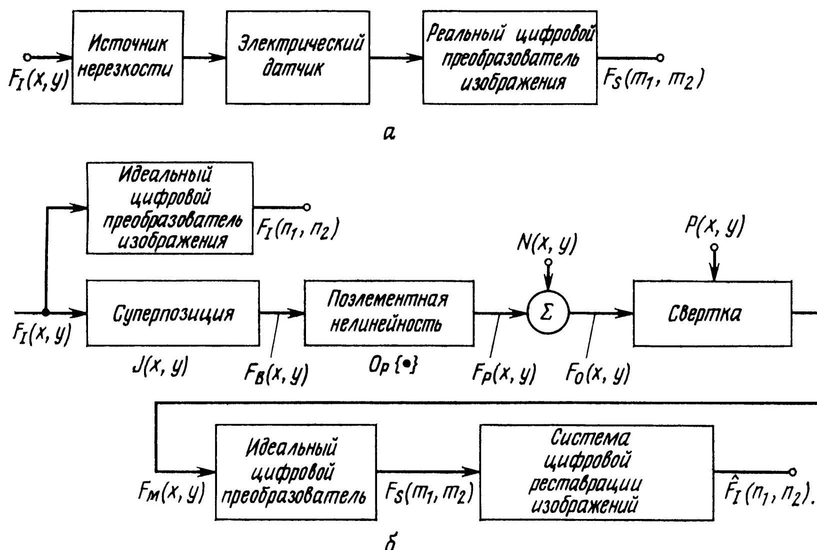


Рис. 13.7.1. Обобщенная модель изображающей системы (*а*) и цифровой реставрации изображений (*б*) для случая дискретизованного нерезкого изображения с аддитивным шумом.

изображения, представленного функцией $F_I(x, y)$, которые образуются на выходе идеального цифрового преобразователя, осуществляющего дискретизацию с шагом ΔI. Для разработки модели цифровой реставрации требуется определить количественные соотношения между отсчетами наблюдаемого изображения $F_S(m_1, m_2)$ и значениями исходного изображения в узловых точках, или псевдоотсчетами, $F_I(n_1, n_2)$, пользуясь методами, изложенными в разд. 9.2. С этой целью проводят усечение эквивалентного импульсного отклика дискретизирующего импульса $P(x, y)$, вводя некоторые пространственные границы $\pm T_P$, а затем в узлах решетки с шагом ΔP выделяют псевдоотсчеты наблюдаемого непрерывного изображения, описываемого функцией $F_O(x, y)$. Следующий шаг в получении дискретного представления — определение соотношений между элементами наблюдаемого поля изображения $F_O(x, y)$ и элементами поля изоб-

ражения $F_P(x, y)$ и шумового поля $N(x, y)$. Последний шаг при разработке модели цифровой реставрации изображений состоит в нахождении дискретного эквивалента операции суперпозиции с использованием импульсного отклика $J(x, y)$. В рассмотренной модели существуют два потенциальных источника погрешностей: усечение импульсных откликов $J(x, y)$ и $P(x, y)$ и интегрирование методом квадратурных сумм. Чтобы можно было пренебречь погрешностями обоих видов, интервалы усечения выбирают большими, а приращения ΔI и ΔP в квадратурных суммах берут малыми. Конечно, при этом увеличиваются размеры массивов и, следовательно, возрастают требования к объему памяти и мощности вычислительных средств. Более того, как будет показано в дальнейшем, повышение точности дискретизации может даже привести к неустойчивости вычислительного процесса реставрации!

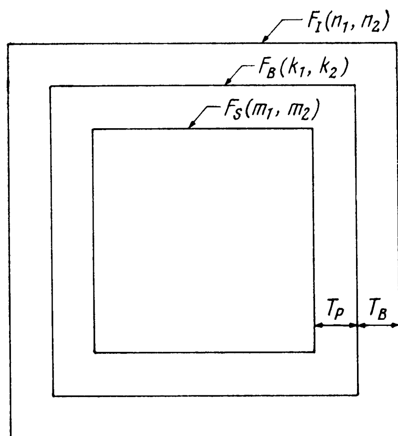


Рис. 13.7.2. Соотношение между различными массивами отсчетов изображения.

Важное значение имеют относительные геометрические размеры различных массивов модели реставрации изображений. Эти массивы образуют гнездо (рис. 13.7.2). Область массива нерезкого изображения $F_B(k_1, k_2)$ меньше области массива идеально дискретизованного изображения $F_I(n_1, n_2)$ на половину протяженности усеченного импульсного отклика $J(x, y)$. Аналогично область массива физических отсчетов $F_S(m_1, m_2)$ меньше области массива наблюдаемых элементов изображения $F_O(k_1, k_2)$ на половину протяженности усеченного импульсного отклика $P(x, y)$.

Чтобы можно было воспользоваться формальной структурой векторной алгебры при последующем анализе процедур реставрации изображений, образуем векторные эквиваленты различных массивов, используемых в реставрационной модели. Следуя методам разд. 9.2, и в этом случае изменим индексы массивов таким образом, чтобы в каждом из них первый элемент занял верхний левый угол.

Затем установим векторные соотношения между отдельными блоками модели, для чего произведем развертку столбцов указанных массивов.

В результате получим

$$\mathbf{f}_S = \mathbf{B}_P \mathbf{f}_O, \quad (13.7.3а)$$

$$\mathbf{f}_O = \mathbf{f}_P + \mathbf{n}, \quad (13.7.3.б)$$

$$\mathbf{f}_P = \mathcal{O}_P \{\mathbf{f}_B\}, \quad (13.7.3в)$$

$$\mathbf{f}_B = \mathbf{B}_B \mathbf{f}_I, \quad (13.7.3г)$$

где $\mathbf{B}_P$ и $\mathbf{B}_B$ — матрицы нерезкости, содержащие отсчеты $P(x, y)$ и $J(x, y)$ соответственно. Нелинейная операция (13.7.3в) определяется как поэлементное нелинейное преобразование. Таким образом, имеем

$$f_P(i) = \mathcal{O}_P \{f_B(i)\}. \quad (13.7.4)$$

Из формул (13.7.3а) и (13.7.3г) можно получить соотношение, описывающее наблюдаемые физические отсчеты изображения, выраженные через математические псевдоотсчеты идеального изображения, следующего вида:

$$\mathbf{f}_S = \mathbf{B}_P \mathcal{O}_P \{\mathbf{B}_B \mathbf{f}_I\} + \mathbf{B}_P \mathbf{n}. \quad (13.7.5)$$

Определим теперь некоторые специальные случаи для этого уравнения. Первый из них — отсутствие поэлементной нелинейности — соответствует равенству

$$\mathbf{f}_S = \mathbf{B} \mathbf{f}_I + \mathbf{n}_B, \quad (13.7.6)$$

где $\mathbf{B} = \mathbf{B}_P \mathbf{B}_B$ и $\mathbf{n}_B = \mathbf{B}_P \mathbf{n}$. Мы получили классическую дискретную модель, содержащую систему линейных уравнений с включением неопределенности измерения. Второй особый случай, который будет рассматриваться позднее, соответствует условию пренебрежимо малой пространственной нерезкости, вносимой цифровым преобразователем. В этом случае можно записать

$$\mathbf{f}_S = \mathcal{O}_P \{\mathbf{B} \mathbf{f}_I\} + \mathbf{n}, \quad (13.7.7)$$

где $\mathbf{B} = \mathbf{B}_B$. Если приращение в квадратурных суммах ΔI взять равным шагу отсчетов ΔS, то матрицы в (13.7.6) и (13.7.7) примут особую форму:

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{1,1} & \mathbf{B}_{1,2} & \cdots & \mathbf{B}_{1,L} & 0 & \cdots & 0 \\ 0 & \mathbf{B}_{2,2} & \cdots & \mathbf{B}_{2,L} & \mathbf{B}_{2,L+1} & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & \mathbf{B}_{M,N-L+1} & & \cdots & \mathbf{B}_{M,N} \end{bmatrix}, \quad (13.7.8)$$

где

$$B_{m_2, n_2}(m_1, n_1) = H(m_1 - n_1 + L, \; m_2 - n_2 + L) \quad (13.7.9)$$

$$\text{при} \quad 1 \leqslant m_i \leqslant M \quad \text{и} \quad m_i \leqslant n_i \leqslant m_i + L - 1.$$

Для проведения экспериментов по реставрации изображений с использованием моделей (13.7.6) и (13.7.7) были сформированы два искусственных изображения [22]. Исследования методов

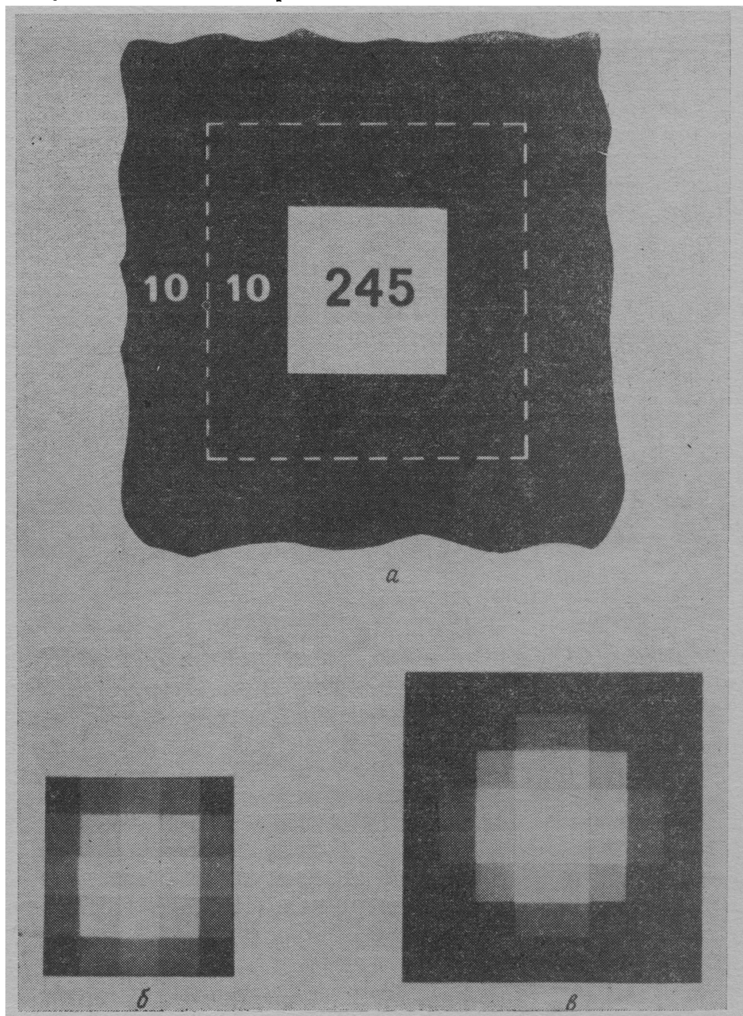


Рис. 13.7.3. Массивы изображения, использованные при исследовании недоопределенной модели.

а — исходное изображение; *б* — импульсный отклик; *в* — наблюдаемое изображение.

реставрации при недоопределении проводились с использованием исходного изображения в виде квадрата размером 4×4 элемента с уровнем 245, находящегося на фоне с уровнем 10

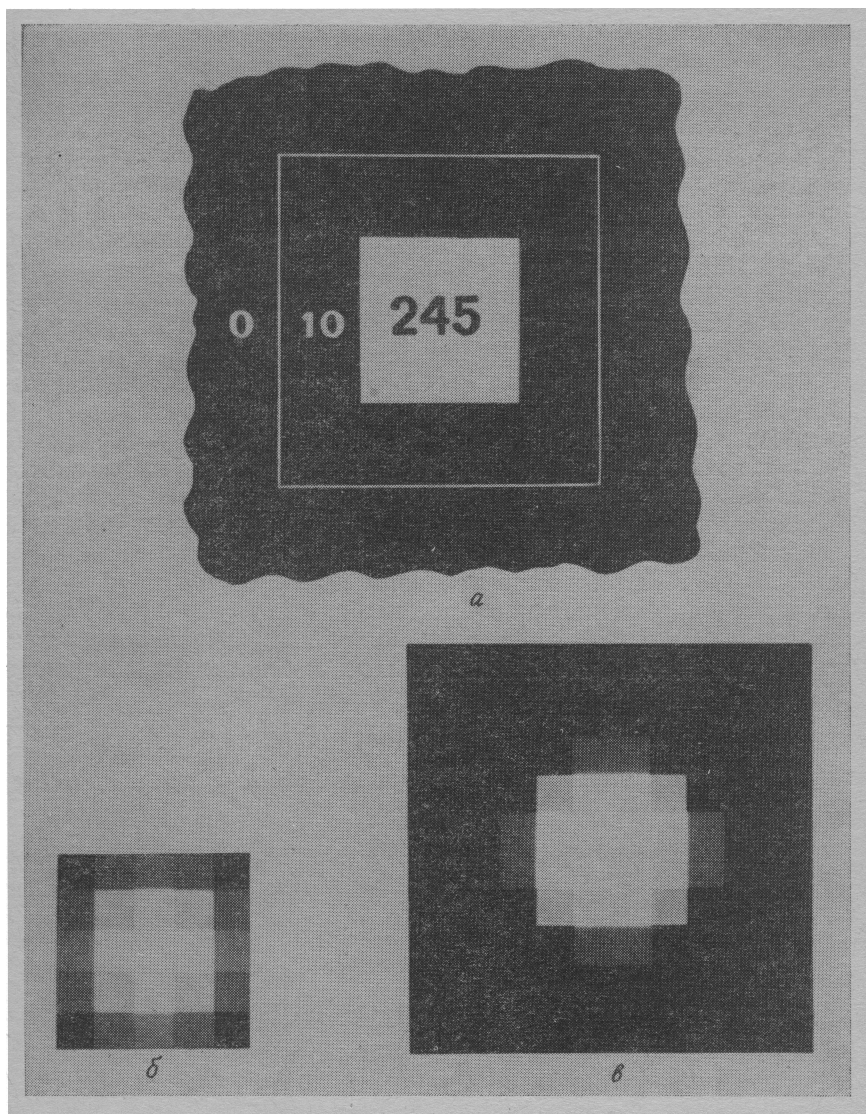


Рис. 13.7.4. Массивы изображения, использованные при исследовании переопределенной модели.
а — исходное изображение; *б* — импульсный отклик; *в* — наблюдаемое изображение.

(рис. 13.7.3, *а*); полная шкала яркостей охватывает уровни от 0 до 255. Функция рассеяния точки гауссовой формы, которую иллюстрирует рис. 13.7.3, *б*, определяется выражением вида

$$H(l_1, l_2) = K \exp\left\{-\left(\frac{l_1^2}{2b_C^2} + \frac{l_2^2}{2b_R^2}\right)\right\} \tag{13.7.10}$$

на массиве размером 5×5 элементов, где K — нормирующий коэффициент, b_C и b_R — коэффициенты пространственной нерезкости. Исходное изображение, которое использовалось для исследования методов реставрации при переопределении (рис. 13.7.4, *а*), имеет вид квадрата размером 4×4 элемента с уровнем 245, расположенного в центре другого квадрата размером 8×8 элементов с уровнем 10 на фоне с нулевой яркостью. Это изображение может служить моделью реального изображения в виде луны на фоне темного неба. Исходное изображение делалось нерезким с помощью импульсного отклика размером 5×5 элементов, показанного на рис. 13.7.4,*б*. Нерезкое изображение состоит из 12×12 элементов. Предполагается, что наблюдаются все 12×12 элементов и что имеется априорное знание об ограниченном размере исходного изображения, состоящего из 8×8 элементов. Из этого следует, что многие элементы матрицы нерезкости (13.7.8) тождественно равны нулю. Более того, обозначив через $\mathbf{f}$ просканированное по столбцам исходное изображение размера 8×8 элементов ($M = 8$), можно предложить следующую векторую модель для 12×12 реально наблюдаемых элементов ($N = 2$):

$$\mathbf{g} = \mathbf{D}\mathbf{f} + \mathbf{n}, \tag{13.7.11}$$

где $\mathbf{D}$ — матрица размера $M^2 \times N^2$

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_{1,1} & 0 & \cdots & 0 \\ \mathbf{D}_{2,1} & \mathbf{D}_{2,2} & & \vdots \\ \cdot & \cdot & & \vdots \\ \cdot & \cdot & & 0 \\ \cdot & \cdot & & \mathbf{D}_{M,N-L+1} \\ \mathbf{D}_{L,1} & \mathbf{D}_{L,2} & & \vdots \\ 0 & \mathbf{D}_{L+1,2} & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & \vdots & \cdots 0 & \mathbf{D}_{M,N} \end{bmatrix}, \tag{13.7.12}$$

причем

$$D_{m_2, n_2}(m_1, n_1) = H(m_1 - n_1 + 1, m_2 - n_2 + 1) \tag{13.7.13}$$

при $1 \leqslant n_i \leqslant N$ и $n_i \leqslant m_i \leqslant n_i + L - 1$. Матрица (13.7.12) представляет собой матрицу оператора конечной свертки D-типа, введенного в гл. 9. Сравнение структур матриц нерезкости (13.7.8) и (13.7.12) для недоопределенной и переопределенной моделей показывает, что эти матрицы являются транспонированными парами. Таким образом, $\mathbf{B} = \mathbf{D}^T$; в особом случае разделимости строк и столбцов можно записать $\mathbf{B}_R = \mathbf{D}_R^T$ и $\mathbf{B}_C = \mathbf{D}_C^T$.

В машинных экспериментах по реставрации изображений наблюдаемые нерезкие изображения для недоопределенной и переопределенной моделей были получены путем умножения развернутых по столбцам матриц исходных изображений (рис. 13.7.3, *а* и 13.7.4, *а*) на матрицы нерезкости **B** и **D** соответственно. Аддитивный белый гауссовый шум, добавляемый к нерезким изображениям, вырабатывался с помощью генератора случайных чисел. Перед воспроизведением все реставрированные изображения подвергались ограничению уровня (поэлементно), чтобы диапазон уровней составил 0—255.

Литература

1. Born M., Wolf E., Principles of Optics, Pergamon Press, New York, 1969 [Имеется перевод: Борн М., Вольф Э., Основы оптики. — М.: Наука, 1970.9]
2. Goodman J. W., Introduction to Fourier Optics, McGraw-Hill, New York, 1968. [Имеется перевод: Гудмен Дж., Введение в Фурье-оптику. — М.: Мир, 1970.]
3. O'Neill E. L., Introduction to Statistical Optics, Addison-Wesley, Reading, Mass., 1963. [Имеется перевод: О'Нейл Э., Введение в статистическую оптику. — М.: Мир, 1966.]
4. Hopkins H. H., *Proc. Roy. Soc.*, *Series A*, **231**, 1184, 98 (July 1955).
5. Hufnagel R. E., Stanley N. R., Modulation Transfer Function Associated with Image Transmission through Turbulent Media, *J. Opt. Soc. Am.*, **54**, 1, 52—61 (January 1964).
6. Henney K., Dudley B., Handbook of Photography, McGraw-Hill, New York, 1939.
7. Evans R. M., Hanson W. T., Brewer W. L., Principles of Color Photography, Wiley, New York, 1953.
8. Mees C. E., The Theory of Photographic Process, MacMillan, New York, 1966.
9. Johnson J. B., Thermal Agitation of Electricity in Conductors, *Phys. Rev.*, **32**, 97—109 (July 1928).
10. Nyquist H., Thermal Agitation of Electric Charge in Conductors, *Phys. Rev.*, **32**, 110—113 (July 1928).
11. Pratt W. K., Laser Communication Systems, Wiley, New York, 1969.
12. Gadsen M., Some Statistical Properties of Pulses from Photomultipliers, *Appl. Opt.*, **4**, 11, 1446—1452 (November 1965).
13. Jamieson J. A. et al., Infrared Physics and Engineering, McGraw-Hill, New York, 1963.
14. Thomas W., Ed., SPSE Handbook of Photographic Science and Engineering, Wiley-Interscience, New York, 1973.
15. Selwyn E. W. H., A Theory of Graininess, *Phot. J.*, **73**, 571 (1935).
16. Selwyn E. W. H., Experiments on the Nature of Graininess, *Phot. J.*, **79**, 513 (1939).

17. Falconer D. G., Noise and Distortion in Photographic Data Storage, *IBM J. Res. Dev.*, 521—526 (September 1970).
18. Huang T. S., Some Notes on Film Grain Noise, Appendix 14 in Restoration of Degraded Images, NSF Summer Study Report, Woods Hole, Mass., 1966, pp. 105—109.
19. Hunt B. R., Digital Image Processing, *Proc. IEEE*, **63**, 4, 693—708 (April 1975). [Имеется перевод: Цифровая обработка изображений. — ТИИЭР, 1975, т. 63, № 4, с. 177—195.]
20. Naderi F., Sawchuk A. A., Nonlinear Detection and Estimation of Images Degraded by Film-Grain Noise, in Proceedings SPIE/OSA Conference on Image Processing, Urbach J. C., Ed., Pacific Grove, California, Vol. **74**, February 1976, pp. 17—24.
21. Kelly D. H., Systems Analysis of Photographic Process-I: A Three Stage Model, *J. Opt Soc. Am.*, **50**, 269 (1960).
22. Mascarenhas N. D. A., Pratt W. K., Digital Image Restoration Under a Regression Model, *IEEE Trans. Circuits and Systems*, CAS-**22**, 3, 252—266 (March 1975).

Глава 14

АЛГЕБРАИЧЕСКИЕ МЕТОДЫ ПРОСТРАНСТВЕННОЙ РЕСТАВРАЦИИ ИЗОБРАЖЕНИЙ

Часто возникают задачи пространственной реставрации изображений с целью устранения геометрических искажений, коррекции нерезкости и снижения уровня шумов различного происхождения. В данной главе описаны алгебраические методы пространственной реставрации изображений. В последующих главах рассматриваются специальные методы пространственной реставрации. Обзорный материал по методам пространственной реставрации содержится в [1—5].

14.1. МЕТОДЫ ПРОСТРАНСТВЕННОЙ ФИЛЬТРАЦИИ НЕПРЕРЫВНЫХ ИЗОБРАЖЕНИЙ

Существует класс изображающих систем, в которых пространственные искажения можно моделировать, пользуясь независимым от линейного сдвига импульсным откликом и аддитивным шумом. В таких системах реставрацию непрерывных изображений можно осуществить методами линейной фильтрации, описываемыми ниже. Эти методы затем обобщаются на случай реставрации дискретных изображений.

14.1.1. ИНВЕРСНАЯ ФИЛЬТРАЦИЯ

Самые ранние попытки осуществить реставрацию изображений были основаны на использовании идеи инверсной фильтрации [6—10]. Анализ инверсной фильтрации можно провести с помощью схемы на рис. 14.1.1. Идеальное изображение $F_I(x, y)$ проходит через линейную искажающую систему с импульсным откликом $H_D(x, y)$ и повреждается аддитивным шумом $N(x, y)$. Предполагается, что между шумом и идеальным изображением корреляция отсутствует. Наблюдаемое изображение можно описать интегралом свертки

$$F_O(x, y) = \int_{-\infty}^{\infty}\int F_I(\alpha, \beta) H_D(x-\alpha, y-\beta)\, d\alpha\, d\beta + N(x, y) \quad (14.1.1а)$$

или

$$F_O(x, y) = F_I(x, y) * H_D(x, y) + N(x, y). \quad (14.1.1б)$$

Система реставрации имеет вид фильтра с независимым от линейного сдвига импульсным откликом $H_R(x, y)$. На выходе этого фильтра получается исправленное изображение, описываемое функцией

$$\hat{F}_I(x, y) = [F_I(x, y) * H_D(x, y) + N(x, y)] * H_R(x, y). \quad (14.1.2)$$

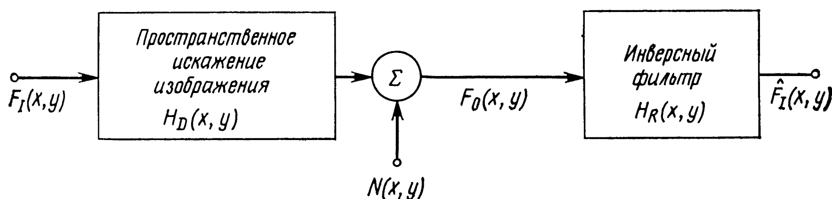


Рис. 14.1.1. Реставрация изображений методом инверсной фильтрации.

Согласно теореме о свертке, имеем

$$\hat{\mathcal{F}}_I(\omega_x, \omega_y) = [\mathcal{F}_I(\omega_x, \omega_y)\mathcal{H}_D(\omega_x, \omega_y) + \mathcal{N}(\omega_x, \omega_y)]\mathcal{H}_R(\omega_x, \omega_y), \quad (14.1.3)$$

где $\mathcal{F}_I(\omega_x, \omega_y)$, $\hat{\mathcal{F}}_I(\omega_x, \omega_y)$, $\mathcal{N}(\omega_x, \omega_y)$, $\mathcal{H}_D(\omega_x, \omega_y)$, $\mathcal{H}_R(\omega_x, \omega_y)$ — двумерные спектры Фурье $F_I(x, y)$, $\hat{F}_I(x, y)$, $N(x, y)$, $H_D(x, y)$, $H_R(x, y)$ соответственно. Если выбрать частотную характеристику $\mathcal{H}_R(\omega_x, \omega_y)$ реставрирующего фильтра так, чтобы

$$\mathcal{H}_R(\omega_x, \omega_y) = 1/\mathcal{H}_D(\omega_x, \omega_y), \quad (14.1.4)$$

то получится спектр исправленного изображения

$$\hat{\mathcal{F}}_I(\omega_x, \omega_y) = \mathcal{F}_I(\omega_x, \omega_y) + [\mathcal{N}(\omega_x, \omega_y)/\mathcal{H}_D(\omega_x, \omega_y)]. \quad (14.1.5)$$

Обратное преобразование Фурье позволяет получить исправленное изображение, описываемое функцией

$$\hat{F}_I(x, y) = F_I(x, y) + (1/4\pi^2)\int_{-\infty}^{\infty}\int \frac{\mathcal{N}(\omega_x, \omega_y)}{\mathcal{H}_D(\omega_x, \omega_y)} \exp\{i[\omega_x x, \omega_y y]\}\, d\omega_x\, d\omega_y. \quad (14.1.6)$$

При отсутствии шума источника достигается идеальная реставрация изображения; в противном случае возникает аддитивная ошибка реставрации, которая может стать очень большой на пространственных частотах с малым значением $\mathcal{H}_D(\omega_x, \omega_y)$. Как правило, $\mathcal{H}_D(\omega_x, \omega_y)$ и $\mathcal{F}_I(\omega_x, \omega_y)$ принимают малые значения в области высоких пространственных частот, поэтому по-

нижение качества изображения выражается в сильном искажении его мелких деталей. Рис. 14.1.2 иллюстрирует частотные зависимости, типичные для инверсной фильтрации.

Шум может существенно затруднить получение однозначной оценки. Другими словами, малые изменения $N(x, y)$ могут стать причиной радикальных изменений оценки $\hat{F}_I(x, y)$. Рассмотрим

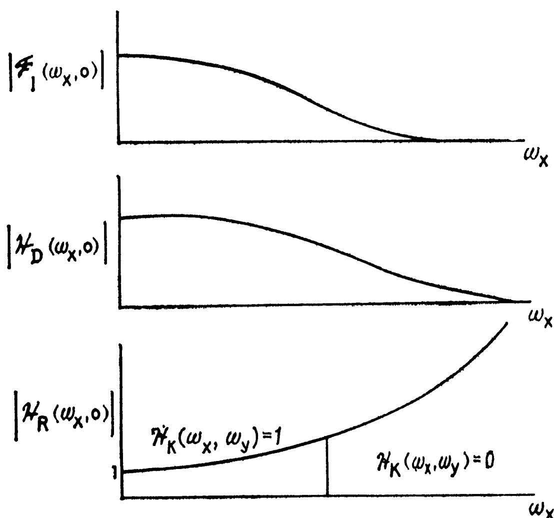


Рис. 14.1.2. Типичные частотные зависимости в системе реставрации изображений методом инверсной фильтрации.

следующий пример. Пусть к функции $F_I(x, y)$, представляющей идеальное изображение, добавляется возмущающая функция $Z(x, y)$. В результате образуется возмущенная функция

$$F_Z(x, y) = F_I(x, y) + Z(x, y). \tag{14.1.7}$$

Существует много возмущающих функций, удовлетворяющих условию

$$\left| \int_{-\infty}^{\infty}\int Z(\alpha, \beta) H_D(x-\alpha, y-\beta)\, d\alpha\, d\beta \right| < |N(x, y)|. \tag{14.1.8}$$

При выполнении этого условия функция $F_Z(x, y)$ может описываться интегралом свертки (14.1.1) в пределах точности измерения наблюдаемого изображения. Можно показать, что в случае использования возмущающей функции в виде высокочастотной синусоиды произвольной амплитуды

$$\lim_{n\to\infty} \left[\int_{-\infty}^{\infty}\int \sin[n(\alpha+\beta)]\, H_D(x-\alpha, y-\beta)\, d\alpha\, d\beta \right] = 0. \tag{14.1.9}$$

С точки зрения задач реставрации этот факт является крайне тревожным по двум причинам. С одной стороны, высокочастотные составляющие, которые могут содержаться в идеальном изображении, будут маскироваться наблюдаемым шумом. С другой стороны, наблюдаемый шум малого уровня может привести к появлению интенсивных высокочастотных составляющих в исправленном изображении, которые отсутствуют в оригинале. Если относительно слабый шум $N(x, y)$ в наблюдаемом изображении соответствует большим значениям возмущающих функций, то говорят, что интегральное уравнение вида (14.1.1а) является неустойчивым или плохо обусловленным. Эта неустойчивость определяется структурой импульсного отклика.

Предложены некоторые частные способы ослабления шумов, возникающих при инверсной фильтрации; один из них [8] заключается в применении реставрирующего фильтра с частотной характеристикой

$$\mathscr{H}_R(\omega_x, \omega_y) = \mathscr{H}_K(\omega_x, \omega_y)/\mathscr{H}_D(\omega_x, \omega_y), \qquad (14.1.10)$$

где $\mathscr{H}_K(\omega_x, \omega_y)$ — функция, принимающая почти единичное значение на тех пространственных частотах, на которых ожидаемая амплитуда спектра идеального изображения превышает ожидаемую амплитуду шумового спектра, и почти нулевое значение на всех остальных пространственных частотах. При этом спектр исправленного изображения имеет вид

$$\hat{\mathscr{F}}_I(\omega_x, \omega_y) = \mathscr{F}_I(\omega_x, \omega_y)\mathscr{H}_K(\omega_x, \omega_y) + \\ + [\mathscr{H}_K(\omega_x, \omega_y)\mathscr{N}(\omega_x, \omega_y)/\mathscr{H}_D(\omega_x, \omega_y)]. \qquad (14.1.11)$$

Этот результат является компромиссом между уменьшением уровня шума и снижением контраста мелких деталей изображения.

Одной из основных трудностей, возникающих при инверсной фильтрации, является также то, что частотная характеристика искажающей системы может иметь нулевые значения в рабочей полосе частот. На таких частотах инверсный фильтр не реализуем, так как его частотная характеристика должна принимать бесконечные значения. Поэтому в этих точках ее аппроксимируют частотной характеристикой, принимающей максимальные значения.

14.1.2. ВИНЕРОВСКАЯ ФИЛЬТРАЦИЯ И АДДИТИВНЫЙ ШУМ

Не следует удивляться тому, что инверсная фильтрация дает плохие результаты при наличии шума, поскольку этот метод совершенно не учитывает шумовых эффектов. Методы винеровской фильтрации, принимающие в расчет априорное знание статистических свойств шумового поля [11—15], позволяют повысить качество исправленных изображений. В последующих разделах

рассматриваются обобщения винеровской фильтрации для случайного идеального изображения и случайного импульсного отклика, описывающего нерезкость. Здесь представлена только модель реставрации, показанная на рис. 14.1.1, когда нерезкое изображение искажается аддитивным шумом. Предполагается,

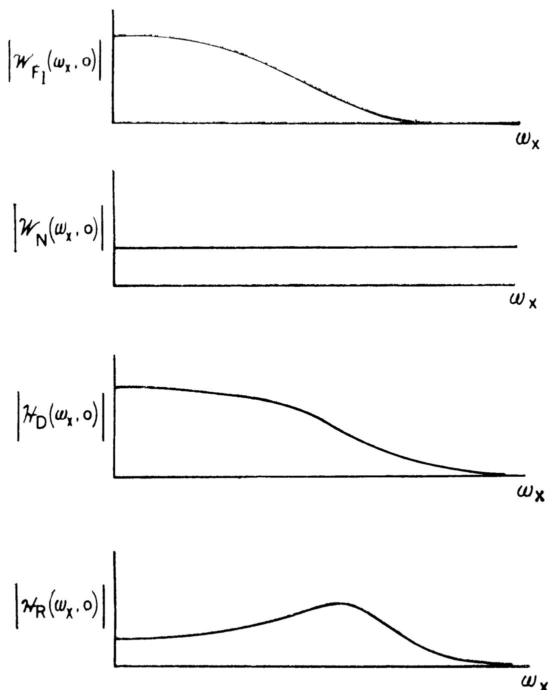


Рис. 14.1.3. Типичные частотные зависимости в системе реставрации изображений методом винеровской фильтрации.

что шум имеет нулевое среднее значение, не зависит от изображения и его энергетический спектр $\mathscr{W}_N(\omega_x, \omega_y)$ известен.

В случае винеровской фильтрации непрерывного изображения импульсный отклик реставрирующего фильтра выбирают таким образом, чтобы минимизировать среднеквадратическую ошибку реставрации

$$\mathscr{E} = E\{[F_I(x, y) - \hat{F}_I(x, y)]^2\}. \qquad (14.1.12)$$

Частотная характеристика фильтра с таким импульсным откликом имеет вид [11]

$$\mathscr{H}_R(\omega_x, \omega_y) = \frac{\mathscr{H}_D^*(\omega_x, \omega_y)}{|\mathscr{H}_D(\omega_x, \omega_y)|^2 + \mathscr{W}_N(\omega_x, \omega_y)}, \qquad (14.1.13)$$

где $\mathscr{H}_D(\omega_x, \omega_y)$ — частотная характеристика искажающей системы [1]. В пределе, когда шум отсутствует, реставрирующий винеровский фильтр становится инверсным фильтром. Рис. 14.1.3 иллюстрирует типичные спектр изображения, частотную характеристику искажающей системы и результирующую частотную характеристику винеровского фильтра. Последняя имеет плавный спад естественного вида, обусловленный учетом энергетического спектра шума.

14.1.3. ВИНЕРОВСКАЯ ФИЛЬТРАЦИЯ И СЛУЧАЙНОЕ ИЗОБРАЖЕНИЕ

Винеровский фильтр с частотной характеристикой (14.1.13), учитывающий аддитивный шум, по существу полностью игнорирует пространственную корреляцию, присущую идеальному изображению. Использование априорных данных о статистических свойствах пространственной корреляции часто позволяет существенно улучшить результат реставрации. Такое обобщение винеровского фильтра основано на предположении, что идеальное изображение является реализацией двумерного случайного процесса с известным энергетическим спектром $\mathscr{W}_{F_I}(\omega_x, \omega_y)$. Для упрощения аналитических выражений также принимается, что процесс, представляющий изображение, имеет нулевое среднее. В рамках всех этих допущений частотная характеристика реставрирующего фильтра, обеспечивающего минимальную среднеквадратическую ошибку, имеет вид

$$\mathscr{H}_R(\omega_x, \omega_y) = \frac{\mathscr{H}_D^*(\omega_x, \omega_y)\,\mathscr{W}_{F_I}(\omega_x, \omega_y)}{|\mathscr{H}_D(\omega_x, \omega_y)|^2\,\mathscr{W}_{F_I}(\omega_x, \omega_y) + \mathscr{W}_N(\omega_x, \omega_y)}, \quad (14.1.14а)$$

или

$$\mathscr{H}_R(\omega_x, \omega_y) = \frac{\mathscr{H}_D^*(\omega_x, \omega_y)}{|\mathscr{H}_D(\omega_x, \omega_y)|^2 + \mathscr{W}_N(\omega_x, \omega_y)/\mathscr{W}_{F_I}(\omega_x, \omega_y)}. \quad (14.1.14б)$$

В формуле (14.1.14б) частотная характеристика реставрирующего фильтра выражена через отношение сигнал/шум (С/Ш) для энергетических спектров

$$\text{С/Ш}(\omega_x, \omega_y) \equiv \mathscr{W}_{F_I}(\omega_x, \omega_y)/\mathscr{W}_N(\omega_x, \omega_y). \quad (14.1.15)$$

Если идеальное изображение не обладает пространственной корреляцией, его энергетический спектр принимает единичное значение во всем диапазоне пространственных частот и соотношение (14.1.14) переходит в (14.1.13).

[1] Полный вывод винеровских оценок дискретных изображений содержится в последующих разделах.

14.1.4. ВИНЕРОВСКАЯ ФИЛЬТРАЦИЯ И СЛУЧАЙНАЯ НЕРЕЗКОСТЬ

Во многих изображающих системах импульсный отклик, описывающий нерезкость, оказывается непостоянным: его форма может изменяться случайным образом. Практический пример — нерезкость, возникающая при наблюдении через турбулентную атмосферу. Очевидно, что винеровский фильтр сможет обеспечить лучшие результаты, если он будет динамически адаптироваться на изменения импульсного отклика нерезкости. Когда это невозможно, следует добиться улучшения винеровского фильтра на этапе его проектирования, считая, что импульсный отклик является реализацией двумерного случайного процесса с известным средним (определяющим «усредненную» форму импульсного отклика) и наложенным на него случайным возмущением, заданным известным энергетическим спектром. Слепян [16] определил частотные характеристики такого реставрирующего фильтра для нескольких мер ошибок.

14.1.5. ФИЛЬТРЫ НА ОСНОВЕ ПАРАМЕТРИЧЕСКИХ ОЦЕНОК

Для осуществления реставрации изображений разработаны различные варианты винеровского фильтра. Одни из них являются эвристическими, другие — основаны на теории.

Стокхэм и Коул [17] предложили реставрирующий фильтр с частотной характеристикой

$$\mathscr{H}_R(\omega_x, \omega_y) = \left[\frac{\mathscr{W}_{F_I}(\omega_x, \omega_y)}{|\mathscr{H}_D(\omega_x, \omega_y)|^2 \mathscr{W}_{F_I}(\omega_x, \omega_y) + \mathscr{W}_N(\omega_x, \omega_y)}\right]^{1/2}. \tag{14.1.16}$$

Энергетический спектр изображения на выходе реставрирующего фильтра имеет вид

$$\mathscr{W}_{\hat{F}_I}(\omega_x, \omega_y) = |\mathscr{H}_R(\omega_x, \omega_y)|^2 \mathscr{W}_{F_O}(\omega_x, \omega_y), \tag{14.1.17}$$

где $\mathscr{W}_{F_O}(\omega_x, \omega_y)$ — энергетический спектр наблюдаемого изображения, связанный с энергетическим спектром идеального изображения следующим соотношением:

$$\mathscr{W}_{F_O}(\omega_x, \omega_y) = |\mathscr{H}_D(\omega_x, \omega_y)|^2 \mathscr{W}_{F_I}(\omega_x, \omega_y) + \mathscr{W}_N(\omega_x, \omega_y). \tag{14.1.18}$$

Нетрудно заметить, что энергетический спектр исправленного изображения совпадает с энергетическим спектром идеального изображения, т. е.

$$\mathscr{W}_{\hat{F}_I}(\omega_x, \omega_y) = \mathscr{W}_{F_I}(\omega_x, \omega_y). \tag{14.1.19}$$

Поэтому реставрирующий фильтр с характеристикой (14.1.16) будем называть фильтром, сохраняющим энергетический спектр изображения. Для сравнения запишем выражение для энергетического спектра исправленного изображения при использовании винеровского фильтра с частотной характеристикой (14.1.14):

$$\mathscr{W}_{\hat{F}_I}(\omega_x, \omega_y) = \frac{|\mathscr{H}_D(\omega_x, \omega_y)|^2 [\mathscr{W}_{F_I}(\omega_x, \omega_y)]^2}{[|\mathscr{H}_D(\omega_x, \omega_y)|^2 \mathscr{W}_{F_I}(\omega_x, \omega_y) + \mathscr{W}_N(\omega_x, \omega_y)]}. \quad (14.1.20)$$

В данном случае исправленное и идеальное изображения будут иметь одинаковые энергетические спектры только при отсутствии шума. Обеспечение совпадения энергетических спектров идеального и исправленного изображений, по-видимому, является привлекательным свойством фильтра с частотной характеристикой (4.1.16). Однако гораздо важнее было бы совпадение спектров Фурье этих изображений, поскольку заданное (не случайное) изображение однозначно определяется спектром Фурье, а энергетический спектр неоднозначно связан с изображением. Кроме того, винеровский фильтр дает оценку с минимальной среднеквадратической ошибкой, а фильтр, сохраняющий энергетический спектр изображения, может привести к большей среднеквадратической ошибке.

Стокхэм и Коул [17] также предложили фильтр более общего типа с частотной характеристикой

$$\mathscr{H}_R(\omega_x, \omega_y) = [\mathscr{H}_D(\omega_x, \omega_y)]^{-S} \times$$
$$\times \left[\frac{\mathscr{H}_D^*(\omega_x, \omega_y) \mathscr{W}_{F_I}(\omega_x, \omega_y)}{|\mathscr{H}_D(\omega_x, \omega_y)|^2 \mathscr{W}_{F_I}(\omega_x, \omega_y) + \mathscr{W}_N(\omega_x, \omega_y)} \right]^{1-S}, \quad (14.1.21)$$

которая определяется параметром S, где $0 \leqslant S \leqslant 1$. При $S = 0{,}5$ и $\mathscr{H}_D = \mathscr{H}_D^*$ этот фильтр становится фильтром, сохраняющим энергетический спектр изображения, с частотной характеристикой (14.1.16).

Хант [18] разработал другой параметрический реставрирующий фильтр — так называемый фильтр обусловленных наименьших квадратов с частотной характеристикой

$$\mathscr{H}_R(\omega_x, \omega_y) = \frac{\mathscr{H}_D^*(\omega_x, \omega_y)}{|\mathscr{H}_D(\omega_x, \omega_y)|^2 + \gamma |\mathscr{E}(\omega_x, \omega_y)|^2}, \quad (14.1.22)$$

где γ — постоянная, $\mathscr{E}(\omega_x, \omega_y)$ — произвольная спектральная функция. Если $\gamma = 1$, а $|\mathscr{E}(\omega_x, \omega_y)|^2$ выбран равным отношению энергетических спектров (14.1.15), то фильтр обусловленных наименьших квадратов становится винеровским с частотной характеристикой (14.1.14б). При другом выборе спектральной функции можно минимизировать производные высшего порядка функции, описывающей исправленное изображение.

14.1.6. ФИЛЬТРАЦИЯ ДИСКРЕТНЫХ ИЗОБРАЖЕНИЙ

Методы фильтрации — инверсной, винеровской и на основе параметрической оценки, разработанные для обработки непрерывных изображений, часто применяют для реставрации дискретных изображений. Общий метод состоит в замене всех непрерывных спектральных функций, используемых при фильтрации, соответствующими двумерными дискретными спектрами Фурье. При выполнении преобразований следует заботиться о том, чтобы операция дискретной фильтрации точно отображала процедуру непрерывной свертки, когда

$$\hat{F}_I(x, y) = F_O(x, y) * H_R(x, y), \tag{14.1.23}$$

и чтобы дискретное представление импульсного отклика реставрирующего фильтра точно моделировало импульсный отклик непрерывного фильтра. Оба этих аспекта обсуждаются ниже.

В соответствии с принципами, изложенными в гл. 11, будем считать, что для представления всех непрерывных функций использованы квадратные решетки с шагом Δ, представляющим собой интервал Найквиста для идеального изображения. Импульсный отклик реставрирующего фильтра после пространственного усечения до $L\Delta$ дискретизируется, в результате чего образуется массив $H_R(j, k)$ размером $L \times L$. Аналогично дискретизация функции $F_O(x, y)$ (описывающей наблюдаемое изображение), усеченной до $N\Delta$, приводит к образованию массива $F_O(j, k)$ размером $N \times N$. Следует отметить, что дискретизация $H_R(x, y)$ является процедурой математического моделирования, а дискретизация наблюдаемого изображения $F_O(x, y)$ — физическим процессом. Дискретная свертка дает массив $\hat{F}_I(j, k)$ размером $M \times M$. Относительные размеры областей, занимаемых этими массивами, иллюстрирует рис. 9.2.3. В результате двумерного дискретного преобразования Фурье массивов $F_O(j, k)$ и $H_R(j, k)$ получаются соответственно выражения

$$\mathcal{F}_O(u, v) = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F_O(j, k) \exp\left\{-\frac{2\pi i}{J}(uj + vk)\right\} \tag{14.1.24}$$

и

$$\mathcal{H}_R(u, v) = \frac{1}{N} \sum_{j=0}^{L-1} \sum_{k=0}^{L-1} H_R(j, k) \exp\left\{-\frac{2\pi i}{J}(uj + vk)\right\}, \tag{14.1.25}$$

где $u, v = 0, 1, 2, \ldots, J-1$. Величина J выбирается таким образом, чтобы $J \geqslant N$ и $J = 2^n$, где n — целое число, позволяющее использовать быстрое преобразование Фурье. Затем спектры Фурье наблюдаемого изображения и импульсного отклика ре-

ставрирующего фильтра скалярно перемножаются, а результат перемножения подвергается обратному преобразованию Фурье. В результате получается массив размером $J \times J$:

$$K_E(j,\ k) = \frac{1}{N} \sum_{u=0}^{J-1} \sum_{v=0}^{J-1} \mathscr{F}_O(u,\ v)\, \mathscr{H}_R(u,\ v) \exp\left\{\frac{2\pi i}{J}(uj + vk)\right\}. \tag{14.1.26}$$

Как видно из рис. 9.4.1, область массива $K_E\ (j,\ k)$ размером $M \times M$ представляет искомое исправленное дискретное изображение $\hat{F}_I\ (j,\ k)$. До сих пор была проведена только одна аппроксимация: пространственное усечение импульсного отклика реставрирующего фильтра. Ошибку усечения можно уменьшить до любой желаемой величины путем увеличения размера массива и объема вычислений. Использование интервала Найквиста обеспечивает принципиальную возможность идеального восстановления функции $\hat{F}_I\ (x,\ y)$ путем интерполяции отсчетов $F_I\ (i,\ k)$, образующих массив размером $M \times M$. Остается решить следующую трудную задачу: определить значения отсчетов импульсного отклика реставрирующего фильтра, используя выражение для частотной характеристики этого фильтра. Теоретически метод решения очевиден. По заданной непрерывной частотной характеристике $\mathscr{H}_R\ (\omega_x,\ \omega_y)$ сначала путем обратного преобразования Фурье находится непрерывный импульсный отклик $H_R\ (x,\ y)$; затем формируется дискретный импульсный отклик $H_R\ (j,\ k)$ путем дискретизации $H_R\ (x,\ y)$ в пределах решетки размером $L \times L$ с шагом дискретизации Δ. На практике, однако, найти аналитически $H_R\ (x,\ y)$ по $\mathscr{H}_R\ (\omega_x,\ \omega_y)$ часто оказывается затруднительным. Альтернативный подход состоит в дискретизации непрерывной функции $\mathscr{H}_R\ (\omega_x,\ \omega_y)$ таким образом, чтобы получить дискретный спектр $\mathscr{H}_R\ (u,\ v)$. Однако и этот метод может привести к аналитическим трудностям в случае использования сложных весовых функций (окон). Если аналитическое преобразование неосуществимо, остается только один выход — вычислить дискретное обратное преобразование Фурье частотной характеристики $\mathscr{H}_R\ (\omega_x,\ \omega_y)$ при $\omega_x = 2\pi u/J\Delta$ и $\omega_y = 2\pi v/J\Delta$. Полученный массив отсчетов спектра импульсного отклика следует обработать соответствующим окном для получения массива размером $L \times L$. В результате дискретного преобразования Фурье ограниченного массива получается искомый массив $\mathscr{H}_R\ (u,\ v)$. Без использования окна исправленное изображение оказывается пораженным циклической ошибкой.

Рассмотренные в данном разделе методы реставрации непрерывных изображений с использованием преобразования Фурье относительно просты для понимания и легко реализуются; в этом

заключается их достоинство. Однако им присущи серьезные ограничения. Во-первых, не принимается никаких мер по борьбе с эффектами ошибок наложения побочных спектров, возникающих при дискретизации наблюдаемого изображения с недостаточно высокой частотой. Во-вторых, в самой формулировке задачи предполагается, что шаг квадратной решетки, используемой для вычисления интеграла свертки, не отличается от периода дискретизации. В-третьих, могут реставрироваться только изображения с линейными пространственно-инвариантными искажениями. В-четвертых (вероятно, это самый главный аргумент), затруднены анализ искажений, вызываемых вычислительными ошибками в процессе реставрации, и развитие методов борьбы с такими ошибками. Все это вынуждает вернуться к дискретной модели в виде массива отсчетов нерезкого изображения, разработанной в разд. 9.2, и сформулировать задачу реставрации изображений как чисто вычислительную. Этим вопросам посвящены остальные разделы данной главы.

14.2. ПРОСТРАНСТВЕННАЯ РЕСТАВРАЦИЯ ИЗОБРАЖЕНИЙ МЕТОДОМ ПСЕВДООБРАЩЕНИЯ МАТРИЦ

Иногда пространственные искажения удается промоделировать некоторой операцией в векторном пространстве над вектором псевдоотсчетов идеального изображения, результатом которой является вектор наблюдаемых отсчетов искаженного изображения; пространственную реставрацию цифровых изображений с таким искажением можно осуществить методом псевдообращения матриц, описанным в гл. 8 [19—21].

14.2.1. РЕСТАВРАЦИЯ НЕРЕЗКИХ ИЗОБРАЖЕНИЙ

Прежде всего применим псевдообращение с целью реставрации нерезкого изображения. Наблюдаемое изображение моделируется интегралом суперпозиции

$$F_O(x, y) = \int_{-\infty}^{\infty} F_I(x, y) J(x, y; \alpha, \beta)\, d\alpha\, d\beta, \qquad (14.2.1)$$

где $F_I(x, y)$ и $F_O(x, y)$ — функции, представляющие идеальное и нерезкое изображения соответственно, $J(x, y; \alpha, \beta)$ — импульсный отклик (который может быть пространственно-зависимым) изображающей системы, вносящей нерезкость. Согласно формуле (13.7.6), дискретный эквивалент соотношения (14.2.1) можно получить в виде матричного уравнения

$$\mathbf{g} = \mathbf{B}\mathbf{f}, \qquad (14.2.2)$$

где $\mathbf{g}$ — вектор размера $P \times 1$ ($P = M^2$), учитывающий $M \times M$ реальных отсчетов нерезкого изображения, $\mathbf{f}$ — вектор размера $Q \times 1$ ($Q = N^2$), учитывающий $N \times N$ псевдоотсчетов идеального изображения, а матрица $\mathbf{B}$ размера $P \times Q$ состоит из элементов, представляющих собой отсчеты импульсного отклика. Если шаг реальных отсчетов совпадает с шагом квадратурных сумм, то $P < Q$, а система уравнений является недоопределенной. Путем увеличения числа отсчетов нерезкого изображения можно обеспечить $P > Q$ и даже $P = Q$. В любом из этих двух случаев систему уравнений называют переопределенной. Систему такого рода можно также получить, если некоторые компоненты вектора идеального изображения выразить через априорные данные. Например, если известно, что идеальное изображение содержит объект конечных размеров на черном фоне (с нулевой яркостью), то компоненты $\mathbf{f}$ вне контура этого объекта можно считать нулевыми.

Задача дискретной реставрации сводится к нахождению решения $\hat{\mathbf{f}}$ уравнения (14.2.2), такого, что

$$\mathbf{B}\hat{\mathbf{f}} = \mathbf{g}. \tag{14.2.3}$$

Поскольку вектор $\mathbf{g}$ задается реальными отсчетами, а элементы матрицы $\mathbf{B}$ определяются независимо моделью системы, не гарантируется даже существование $\hat{\mathbf{f}}$, удовлетворяющего условию (14.2.3). Если система уравнений имеет решение, она называется совместной, если не имеет — несовместной.

В гл. 8. было показано, что несовместность системы уравнений (14.2.2) можно выразить в виде

$$\mathbf{g} = \mathbf{B}\mathbf{f} + \mathbf{e}(\mathbf{f}), \tag{14.2.4}$$

где $\mathbf{e}(\mathbf{f})$ — вектор ошибки, величина которой зависит от $\mathbf{f}$. При несовместной системе уравнений ищется решение

$$\hat{\mathbf{f}} = \mathbf{W}\mathbf{g}, \tag{14.2.5}$$

где $\mathbf{W}$ — линейный оператор, минимизирующий ошибку по критерию наименьших квадратов

$$\mathscr{E}_M = [\mathbf{e}(\hat{\mathbf{f}})]^T [\mathbf{e}(\hat{\mathbf{f}})] = [\mathbf{g} - \mathbf{B}\hat{\mathbf{f}}]^T [\mathbf{g} - \mathbf{B}\hat{\mathbf{f}}]. \tag{14.2.6}$$

В гл. 8 было показано, что эта ошибка минимальна, когда оператор $\mathbf{W}$ выбирается равным $\mathbf{B}^S$ — обратной матрице для $\mathbf{B}$, полученной методом обращения по критерию наименьших квадратов. Такая обратная матрица не обязательно является единственной. В той же главе доказывается, что оператор обобщенного обращения $\mathbf{W} = \mathbf{B}^-$ (особый случай обращения по критерию наименьших квадратов) является единственным, минимизирует ошибку по критерию наименьших квадратов и одновременно обеспечивает

получение оценки с минимальной нормой. Таким образом, сумма квадратов $\hat{\mathbf{f}}$ минимальна при всех возможных оценках по критерию наименьших квадратов. Компенсация нерезкости методом обобщенного обращения матриц обеспечивает получение наименьшей яркости исправленного изображения.

Если система уравнений (14.2.2) является совместной, то условию (14.2.3) может удовлетворять одно или несколько решений. Обычно выбирается оценка, минимизирующая ошибку по критерию наименьших квадратов, которая определяется эквивалентными выражениями вида

$$\mathscr{E}_E = (\mathbf{f} - \hat{\mathbf{f}})^T (\mathbf{f} - \hat{\mathbf{f}}), \qquad (14.2.7а)$$

$$\mathscr{E}_E = \operatorname{tr} \{(\mathbf{f} - \hat{\mathbf{f}}) (\mathbf{f} - \hat{\mathbf{f}})^T\}. \qquad (14.2.7б)$$

Доказано, что ошибка оценивания минимальна, если оценка получена методом обобщенного обращения матриц ($\mathbf{W} = \mathbf{B}^-$). В этом случае результирующая остаточная ошибка оценивания равна

$$\mathscr{E}_E = \mathbf{f}^T [\mathbf{I} - (\mathbf{B}^-\mathbf{B})] \mathbf{f} \qquad (14.2.8а)$$

или

$$\mathscr{E}_E = \operatorname{tr} \{\mathbf{f}\mathbf{f}^T [\mathbf{I} - \mathbf{B}^-\mathbf{B}]\}. \qquad (14.2.8б)$$

При $\mathbf{B}^-\mathbf{B} = \mathbf{I}$, естественно, получаем идеальную оценку.

Таким образом, обобщенное обращение матриц дает оптимальное решение как для совместных, так и для несовместных систем уравнений. Как следует из формулы (8.3.5), обобщенная обратная матрица может быть представлена в виде

$$\mathbf{B}^- = (\mathbf{B}^T\mathbf{B})^{-1} \mathbf{B}^T, \qquad (14.2.9а)$$

если матрица $\mathbf{B}$ размера $P \times Q$ имеет ранг Q. Если матрица $\mathbf{B}$ имеет ранг P, то

$$\mathbf{B}^- = \mathbf{B}^T (\mathbf{B}\mathbf{B}^T)^{-1}. \qquad (14.2.9б)$$

В случае совместной системы уравнений и ранга Q обобщенной обратной матрицы получаем оценку

$$\hat{\mathbf{f}} = \mathbf{B}^-\mathbf{g} = \mathbf{B}^-\mathbf{B}\mathbf{f} = [(\mathbf{B}^T\mathbf{B})^{-1} \mathbf{B}^T] \mathbf{B}\mathbf{f} = \mathbf{f}, \qquad (14.2.10)$$

которая, очевидно, является идеальной. Однако во всех других случаях может появиться остаточная ошибка оценивания. Очевидно, что для получения идеальной оценки желательно иметь дело с переопределенной матрицей нерезкости ранга Q. К сожалению, этот случай редко имеет место при реставрации изображений. Увеличение числа отсчетов нерезкого изображения позволяет получить переопределенную систему уравнений ($P > Q$), однако ранг матрицы нерезкости чаще всего оказывается гораздо

меньше Q. Это объясняется тем, что с увеличением плотности отсчетов возрастает линейная зависимость между строками матрицы нерезкости. Единственный практический случай использования ранга Q обобщенной обратной матрицы соответствует объекту реставрации с ограниченной и известной протяженностью.

Применение обобщенного обращения матриц наталкивается на серьезные трудности, связанные с их размером. Обобщенная обратная матрица — это матрица размера $Q \times P$, где Q — число отсчетов изображения, подлежащих оцениванию, а P — число наблюдаемых отсчетов. При использовании оператора обобщенного обращения (14.2.9) в случае изображений большого размера возникают трудности, связанные с обеспечением достоверности вычисления обобщенных обратных матриц и очень большим числом векторных перемножений в (14.2.5). Если матрица нерезкости $\mathbf{B}$ разделима, т. е.

$$\mathbf{B} = \mathbf{B}_C \otimes \mathbf{B}_R, \tag{14.2.11}$$

где $\mathbf{B}_C$ и $\mathbf{B}_R$ — операторы «строчной» и столбцовой» нерезкости, то можно уменьшить объем вычислений. В этом случае обобщенная матрица разделима в том смысле, что

$$\mathbf{B}^- = \mathbf{B}_C^- \otimes \mathbf{B}_R^-, \tag{14.2.12}$$

где $\mathbf{B}_C^-$ и $\mathbf{B}_R^-$ — обобщенные обратные матрицы для матриц $\mathbf{B}_C$ и $\mathbf{B}_R$. Таким образом, в случае разделимой матрицы нерезкости оценку изображения можно получить, последовательно действуя на все строки массива наблюдаемого изображения обобщенной обратной матрицей строчной нерезкости, а затем обрабатывая каждый столбец массива обобщенной обратной матрицей столбцовой нерезкости.

Реставрацию методом псевдообращения изображений большого размера можно осуществить приближенно. При этом нерезкое изображение разбивается на малые фрагменты, реставрируемые независимо. Процедура пофрагментной реставрации изображений подобна процедуре пофрагментной фильтрации (см. разд. 11.3). Полезно предусматривать перекрытие фрагментов и использовать только центральные отсчеты в каждом реставрированном фрагменте, поскольку именно эти отсчеты обладают наименьшей неопределенностью. В разд. 14.3 описывается эффективный вычислительный алгоритм реставрации изображений методом псевдообращения матриц в случае пространственно-инвариантной нерезкости.

На рис. 14.2.1 показаны нерезкие и исправленные изображения, иллюстрирующие процедуру реставрации методом обобщенного обращения матриц, причем были использованы модели искаженных изображений, представленные на рис. 13.7.3 и 13.7.4. Наблюдаемые изображения не были повреждены шумом. Импульс-

ный отклик нерезкости имел гауссову форму (13.7.10) с $b_R = b_C = 1{,}2$. На рисунке, соответствующем случаю переопределенной модели, воспроизведена только центральная реставриро-

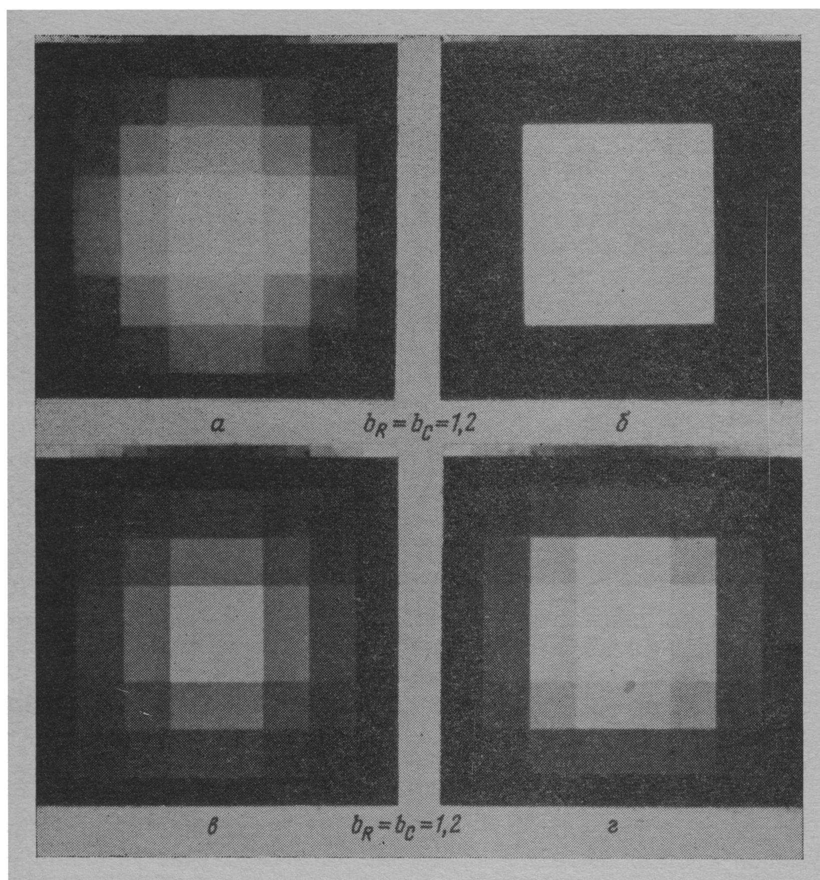


Рис. 14.2.1. Образцы испытательных изображений, исправленных методом псевдообращения матриц. Нерезкость внесена с помощью импульсного отклика гауссовой формы. Наблюдаемое изображение без шума.

а — случай переопределенной модели ($M = 12$, $N = 8$, $L = 5$), нерезкое изображение; *б* — исправленное изображение *а*; *в* — случай недоопределенной модели ($M = 8$, $N = 12$, $L = 5$), нерезкое изображение, минимальная среднеквадратическая ошибка (МСКО) = 4,97 %; *г* — исправленное изображение *в* (МСКО = 1,41 %).

ванная область размером 8×8 элементов, увеличенная до размера 256×256 элементов; нерезкое изображение имеет размер 12×12 элементов. Исправленное изображение визуально не отличается от исходного изображения, однако изучение полученных

числовых данных указывает на наличие вычислительных ошибок округления. В случае недоопределенной модели также воспроизведена только центральная область оценки размером 8×8 элементов. Хотя исправленное изображение выглядит как улучшенное нерезкое изображение, оно не идентично исходному изображению.

14.2.2. РЕСТАВРАЦИЯ НЕРЕЗКОГО ИЗОБРАЖЕНИЯ МЕТОДОМ ПСЕВДООБРАЩЕНИЯ МАТРИЦ ПРИ НАЛИЧИИ ШУМА

Во многих изображающих системах идеальное изображение становится нерезким и, кроме того, повреждается аддитивным шумом. Связь наблюдаемого изображения с идеальным в этом случае можно записать в векторно-матричной форме

$$\mathbf{g} = \mathbf{B}\mathbf{f} + \mathbf{n}, \tag{14.2.13}$$

где $\mathbf{g}$ и $\mathbf{n}$ — векторы размера $P\times 1$, представляющие наблюдаемое изображение и шумовое поле соответственно, $\mathbf{f}$ — вектор размера $Q\times 1$ псевдоотсчетов идеального изображения, $\mathbf{B}$ — матрица нерезкости размера $P\times Q$. Вектор $\mathbf{n}$ состоит из двух аддитивных составляющих, одна из которых включает отсчеты аддитивного внешнего шумового процесса, другая — компоненты разности векторов $(\mathbf{g} - \mathbf{B}\mathbf{f})$, обусловленной ошибками моделирования при задании $\mathbf{B}$. При наличии шума может и не оказаться векторных решений $\hat{\mathbf{f}}$ системы уравнений (14.2.13). Однако (как было показано в гл. 8) обобщенная обратная матрица $\mathbf{B}^-$ позволяет определить оценку по критерию наименьших квадратов и с минимальной нормой. Если нет ошибок моделирования, то оценка

$$\hat{\mathbf{f}} = \mathbf{B}^-\mathbf{g} = \mathbf{B}^-\mathbf{B}\mathbf{f} + \mathbf{B}^-\mathbf{n} \tag{14.2.14}$$

отличается от идеального изображения наличием аддитивной шумовой составляющей $\mathbf{B}^-\mathbf{n}$. Кроме того, в случае недоопределенной модели матрица $\mathbf{B}^-\mathbf{B}$ не является единичной. Ниже рассматривается влияние шума в переопределенной и недоопределенной моделях реставрации изображений.

Переопределенная модель

Если $\mathbf{B}$ есть матрица переопределенной системы уравнений, имеющая ранг Q и заданная выражением (14.2.9а), то $\mathbf{B}^-\mathbf{B} = \mathbf{I}$, а получаемая оценка равна сумме вектора исходного изображения $\mathbf{f}$ и вектора возмущений $\Delta\mathbf{f} = \mathbf{B}^-\mathbf{n}$. Ошибку возмущений, которая содержится в оценке, можно измерить как отношение

нормы вектора возмущений к норме вектора оценки. Можно показать [22, стр. 52], что относительная ошибка ограничена:

$$\|\Delta \mathbf{f}\|/\|\mathbf{f}\| < \|\mathbf{B}^{-}\| \cdot \|\mathbf{B}\| (\|\mathbf{n}\|/\|\mathbf{g}\|). \tag{14.2.15}$$

Произведение $\|\mathbf{B}^{-}\| \cdot \|\mathbf{B}\|$, называемое условным числом $C\{\mathbf{B}\}$ матрицы $\mathbf{B}$, определяет относительную ошибку в оценке, выраженную через отношение норм векторов шумового и наблюдаемого полей. Условное число можно вычислить непосредственно или же определить с помощью отношения наибольшего (W_1) и наименьшего (W_N) сингулярных значений $\mathbf{B}$. Чем больше условное число матрицы нерезкости, тем, очевидно, выше чувствительность к шумовым возмущениям.

На рис. 14.2.2 представлены образцы исправленных изображений для случая переопределенной модели (см. рис. 13.7.4) и гауссовой формы функции нерезкости при нескольких значениях среднеквадратического отклонения и дисперсии шума, равной 10 уровням при полной шкале 0—255. Как и ожидалось, шум наблюдаемого изображения ухудшает качество реставрации. Не было неожиданностью и то, что качество исправленного изображения хуже при средней степени нерезкости и лучше при пониженной нерезкости. Однако затем эта тенденция нарушается: субъективное качество исправленного изображения при очень большой нерезкости выше, чем при нерезкости средней степени. Это кажущееся аномальным поведение, возникающее вследствие пространственного усечения функции рассеяния точки (ФРТ), можно объяснить, наблюдая за условным числом матрицы нерезкости. На рис. 14.2.3. показана зависимость условного числа матрицы нерезкости для рассмотренных примеров от величины коэффициента нерезкости [19]. При небольшой нерезкости условное число мало. Максимум кривой достигается при умеренной нерезкости, а затем следует спад. В области бесконечно больших значений коэффициента нерезкости кривая стремится к горизонтали. Характер кривой объясняет полученные экспериментальные результаты. При осуществлении реставрации производится пространственное усечение ФРТ границами квадратной области, содержащей 5 × 5 псевдоотсчетов — узлов квадратурных сумм. По мере увеличения коэффициента нерезкости при фиксированных M и N ФРТ расширяется и ее хвосты усекаются все сильнее и сильнее. В пределе ненулевые элементы матрицы нерезкости становятся постоянными величинами, а условное число устанавливается на постоянном уровне. При небольших коэффициентах нерезкости эффект усечения незначителен, а кривая условного числа устремляется в бесконечность с асимптотическим значением, полученным в случае неусеченной ФРТ. С увеличением нерезкости число ненулевых элементов матрицы нерезкости возрастает,

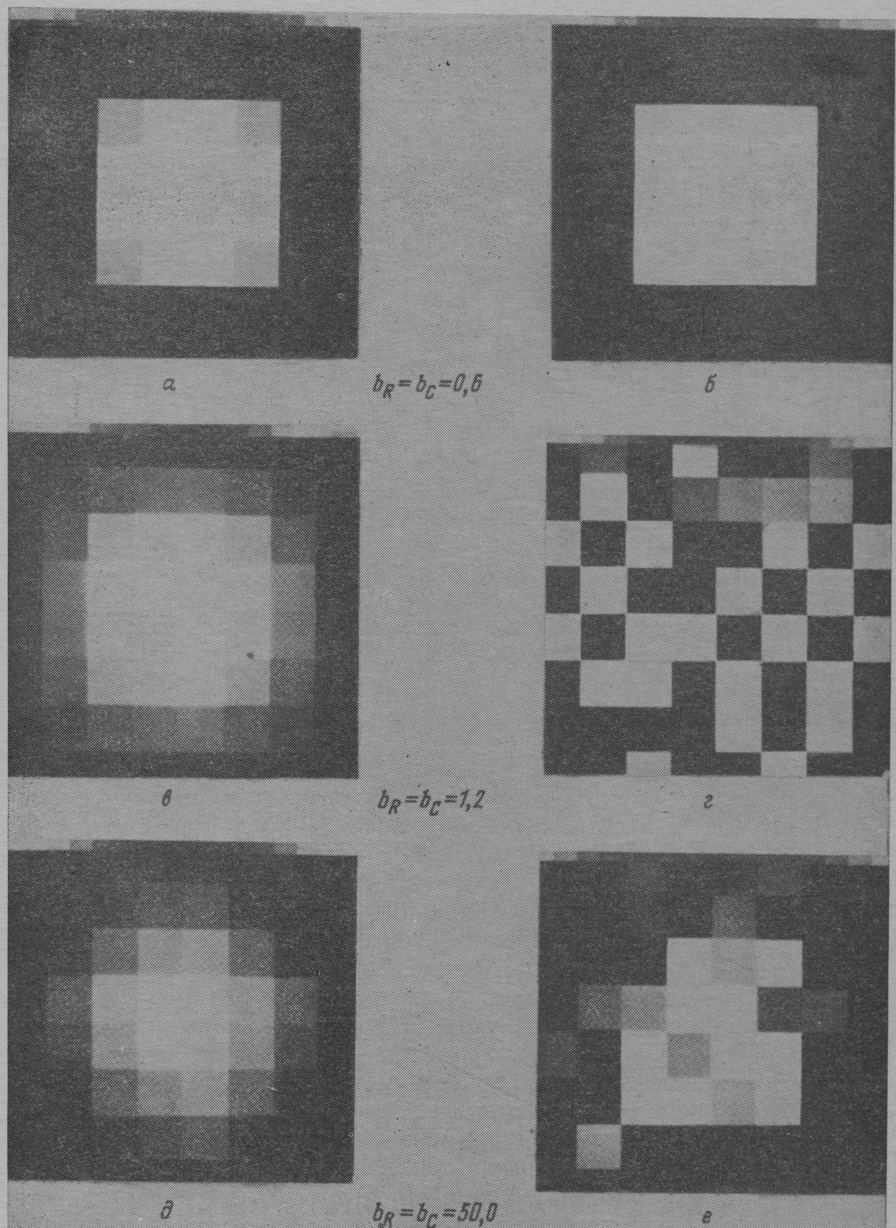


Рис. 14.2.2. Образцы испытательных изображений, исправленных методом псевдообращения матриц. Нерезкость внесена с помощью импульсного отклика гауссовой формы. Случай переопределенной модели ($M = 12$, $N = 8$, $L = 5$). Наблюдаемое изображение с шумом (дисперсия равна 10 уровням шкалы).

а, в, д — нерезкие изображения; *б, г, е* — исправленные изображения.

а условное число стремится к постоянной величине. Таким образом, существует компромисс между числовыми погрешностями вследствие плохой обусловленности и неточностью моделирования. Хотя этот вывод сделан на основе частной модели искажений, он носит, как представляется, более общий характер, поскольку оператор, обратный относительно оператора, описывающего нерезкость, не ограничен. Следовательно, чем точнее дискретная модель аппроксимирует непрерывную модель, тем хуже обусловленность. Более грубая апрроксимация приводит к уменьшению сингулярных значений и увеличению ошибок моделирования. Это противоречие неизбежно, и разрешить его можно только путем использования правильной априорной информации об исходном изображении.

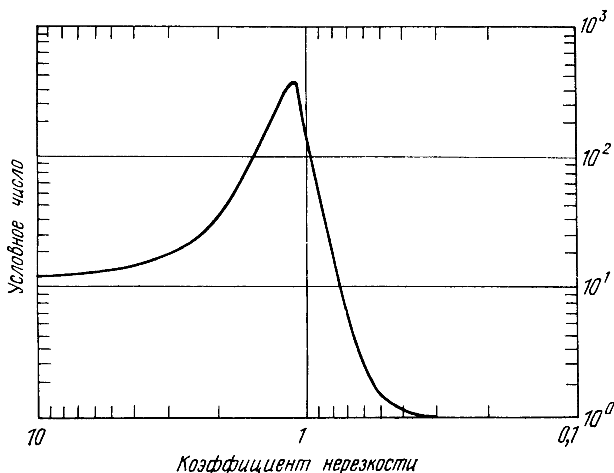


Рис. 14.2.3. Кривая условных чисел.

Недоопределенная модель

На рис. 14.2.4 представлены образцы исправленных изображений в случае недоопределенной модели (см. рис. 13.7.3) и гауссовой нерезкости при нескольких значениях среднеквадратического отклонения и дисперсии шума, равной 10 уровням квантования. Полученные исправленные изображения следуют той же тенденции, которая была отмечена в случае переопределенной модели: наибольшие возмущения в решении появляются при небольших значениях коэффициентов нерезкости. Это нельзя объяснить поведением условного числа, поскольку в случае недоопределен-

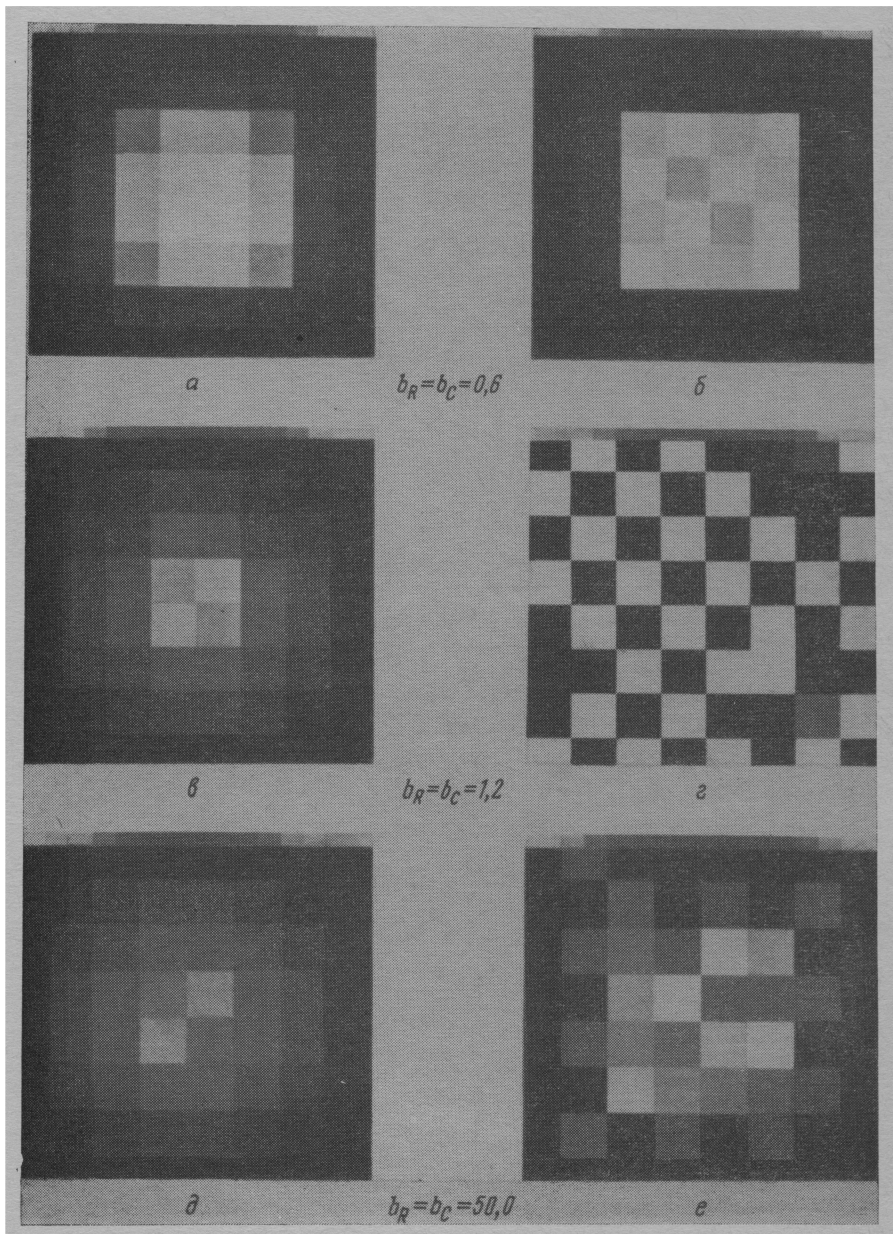


Рис. 14.2.4. Образцы испытательных изображений, исправленных методом псевдообращения матриц. Нерезкость внесена с помощью импульсного отклика гауссовой формы. Случай недоопределенной модели ($M = 8$, $N = 12$, $L = 5$). Наблюдаемое изображение с шумом (дисперсия равна 10 уровням шкалы квантования).

а — нерезкое изображение (МСКО = 1,30 %), *б* — исправленное изображение *а* (МСКО = 0,21 %); *в* — нерезкое изображение (МСКО = 4,91 %), *г* — исправленное изображение *в* (МСКО = 2695,81 %); *д* — нерезкое изображение (МСКО = 7,99 %), *е* — исправленное изображение *д* (МСКО = 7,29 %).

ной модели ($W_N = 0$) оно бесконечно. Однако необходимо учесть, что матрица нерезкости $\mathbf{B}$ получена транспонированием матрицы переопределенной модели и что собственные значения $\mathbf{B}^T\mathbf{B}$ и $\mathbf{B}\mathbf{B}^T$ одинаковы [23, стр. 41]. Следовательно, отношение наибольшего и наименьшего собственных значений $\mathbf{B}^T\mathbf{B}$ определяется кривой на рис. 14.2.3.

14.3. ВЫЧИСЛИТЕЛЬНЫЕ АЛГОРИТМЫ ДЛЯ РЕСТАВРАЦИИ ИЗОБРАЖЕНИЙ МЕТОДОМ ПСЕВДООБРАЩЕНИЯ МАТРИЦ

Прэтт и Давариан [20] разработали эффективные вычислительные алгоритмы для реставрации изображений методом псевдообращения матриц в случае пространственно-инвариантной нерезкости. Эти алгоритмы рассмотрены ниже применительно к переопределенной и недоопределенной моделям изображающей системы.

14.3.1. АЛГОРИТМ ДЛЯ ПЕРЕОПРЕДЕЛЕННОЙ МОДЕЛИ

Вычислительный алгоритм для реставрации методом псевдообращения матриц, разработанный для переопределенной модели, основан на циклической свертке в векторном пространстве (9.3.4), когда

$$\mathbf{k}_E = \mathbf{C}\mathbf{f}_E, \tag{14.3.1}$$

где $\mathbf{C}$ — матрица двумерной циклической свертки, $\mathbf{f}_E$ — вектор изображения размера $J^2 \times 1$, полученный путем введения нулей в вектор $\mathbf{f}$ размера $N^2 \times 1$ идеального изображения, $\mathbf{k}_E$ — соответствующий вектор размера $J^2 \times 1$ выходного изображения. Оценка с использованием обобщенной обратной матрицы записывается непосредственно как

$$\hat{\mathbf{f}}_E = \mathbf{C}^{-1}\mathbf{k}_E, \tag{14.3.2}$$

где $\mathbf{C}^{-1}$ — обратная матрица для матрицы $\mathbf{C}$. Из формул (9.4.4а) и (9.4.5б) легко получить

$$\hat{\mathbf{f}} = [\mathbf{S1}_J^{(N)} \otimes \mathbf{S1}_J^{(N)}]\, \mathbf{C}^{-1} [\mathbf{S1}^{(M)} \otimes \mathbf{S1}_J^{(M)}]^T \mathbf{g}, \tag{14.3.3}$$

где $\mathbf{g} = \mathbf{B}\mathbf{f}$ соответствует случаю переопределенной модели, $\mathbf{S1}^{(K)}$ — выделяющая матрица (9.4.1). Подстановка в выражение (14.3.3) оператора преобразования Фурье $\mathbf{A}_J$ дает

$$\hat{\mathbf{f}} = [\mathbf{S1}_J^{(N)} \otimes \mathbf{S1}_J^{(N)}]\, [\mathbf{A}_J^{-1} \otimes \mathbf{A}_J^{-1}]\, \mathscr{C}^{-1} [\mathbf{A}_J \otimes \mathbf{A}_J]\, [\mathbf{S1}_J^{(M)} \otimes \mathbf{S1}_J^{(M)}]\, \mathbf{g}, \tag{14.3.4}$$

где $\mathscr{C}$ — матрица фильтра для спектрального вычисления свертки, определенная формулой (11.2.13).

Соотношения (14.3.4) и (11.2.13) составляют основу следующего эффективного вычислительного алгоритма реставрации изображений методом псевдообращения матриц в случае переопределенной модели:

а) Поместить матрицу импульсного отклика **H** в верхний левый угол нулевой матрицы размера $J \times J$ для получения расширенной матрицы импульсного отклика ($J \geqslant M$) и подвергнуть эту расширенную матрицу двумерному преобразованию Фурье, что дает

$$\mathscr{H}_E = \mathbf{A}_J \mathbf{H}_E \mathbf{A}_J. \tag{14.3.5}$$

б) Поместить матрицу наблюдаемого изображения **G** в верхний левый угол нулевой матрицы размера $J \times J$ для получения матрицы $\mathbf{K}_E$ и подвергнуть полученную матрицу двумерному преобразованию Фурье, что дает

$$\mathscr{K}_E = \mathbf{A}_J \mathbf{K}_E \mathbf{A}_J. \tag{14.3.6}$$

в) Произвести скалярное деление

$$\hat{\mathscr{F}}_E(m,\ n) = \mathscr{K}_E(m,\ n)/J\mathscr{H}_E(m,\ n), \tag{14.3.7}$$

где $1 \leqslant m \leqslant J$ и $1 \leqslant n \leqslant J$.

г) Вычислить обратное преобразование Фурье

$$\hat{\mathbf{F}}_E = \mathbf{A}_J^{-1} \hat{\mathscr{F}}_E \mathbf{A}_J^{-1}. \tag{14.3.8}$$

д) Выделить искомую матрицу выходного изображения

$$\hat{\mathbf{F}} = [\mathbf{S}1_J^{(N)}]\, \hat{\mathbf{F}}_E\, [\mathbf{S}1_J^{(N)}]^T. \tag{14.3.9}$$

При использовании этого алгоритма требуется выполнить порядка $J^2\ (1 + 4 \log_2 J)$ операций, чтобы вычислить псевдообращенную матрицу методом конечной свертки. Прямое вычисление потребовало бы около M^2N^2 операций. При $M = 256$ и $L = 17$ число операций уменьшается почти в 1750 раз!

Следует отметить, что алгоритм дискретного псевдообращения матриц обходит одну из главных проблем инверсной фильтрации непрерывного изображения: деление на нулевые значения непрерывной частотной характеристики. Ни один элемент матрицы дискретной частотной характеристики, определенной формулой (14.3.5), не может тождественно равняться нулю, если размер L двумерного импульсного отклика представляет собой нечетное целое число, а размер J обрабатываемого массива является четным числом. В предельном случае, когда импульсный отклик **H**

постоянный, получается дискретная частотная характеристика вида

$$\mathcal{H}_E(u, v) = \frac{1}{J} \exp\left\{\frac{-i\pi(L-1)}{J}[u+v]\right\} \frac{\sin[\pi Lu/J]\sin[\pi Lv/J]}{\sin[\pi u/J]\sin[\pi v/J]}, \tag{14.3.10}$$

при $0 \leqslant u,\ v \leqslant J-1$. Эта функция принимает нулевые значения только на нецелочисленных пространственных частотах $u = J/L$ и $v = J/L$, поскольку J четно, а L нечетно.

14.3.2. АЛГОРИТМ ДЛЯ НЕДООПРЕДЕЛЕННОЙ МОДЕЛИ

К сожалению, для недоопределенной модели не существует аналога соотношения (14.3.4), описывающего псевдообращение, поскольку, как пояснено в разд. 9.4, $\hat{\mathbf{f}}$ невозможно выделить из $\mathbf{k}_E$, если $\mathbf{g} = \mathbf{Bf}$ представляет недоопределенную систему уравнений. Однако можно видоизменить наблюдаемое нерезкое изображение $\mathbf{g}$ и благодаря этому сократить объем вычислений, что возможно в случае переопределенной модели. Для упрощения пояснений первоначально ограничимся рассмотрением одномерного примера.

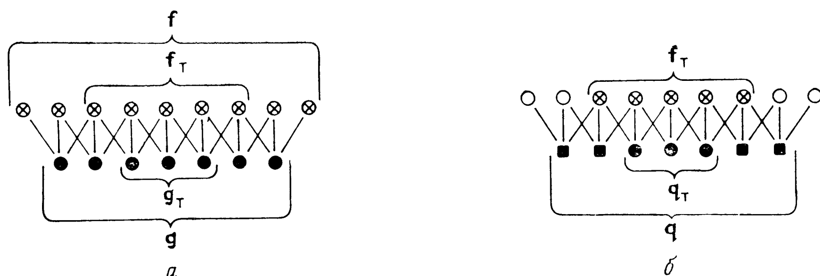


Рис. 14.3.1. Примеры вычисления одномерной свертки.
а — непрерывная свертка с дискретизацией; *б* — дискретная свертка

Пусть вектор $\mathbf{f}_T$ размера $N \times 1$ и вектор $\mathbf{g}_T$ размера $M \times 1$ сформированы из векторов $\mathbf{f}$ и $\mathbf{g}$ отбрасыванием L—1 крайних элементов соответствующего вектора. Можно сказать, что усеченный вектор есть результат действия оператора выделения элементов, т. е.

$$\mathbf{f}_T = \mathbf{S2}_N^{(K)}\mathbf{f}, \tag{14.3.11}$$

где $K = N - 2(L-1)$, и

$$\mathbf{g}_T = \mathbf{S2}_M^{(R)}\mathbf{g}, \tag{14.3.12}$$

где $R = M - 2\ (L-1)$. Рис. 14.3.1, *а* иллюстрирует соотношение между рассматриваемыми векторами для $N = 9$ (число

элементов исходного вектора), $M = 7$ (число элементов наблюдаемого вектора) и $L = 3$ (число элементов импульсного отклика).

Из векторов $\mathbf{f}_T$ и $\mathbf{g}_T$ получаются расширенные векторы $\mathbf{f}_E$ и $\mathbf{g}_E$, связанные следующим матричным уравнением:

$$\mathbf{q}_E = \mathbf{C}\mathbf{f}_E + \mathbf{n}_E, \qquad (14.3.13а)$$

или

$$\begin{matrix} M \{ \\ J-M \{ \end{matrix} \begin{bmatrix} \mathbf{q} \\ \cdots \\ 0 \end{bmatrix} = [\mathbf{C}] \begin{bmatrix} \mathbf{f}_T \\ \cdots \\ 0 \end{bmatrix} \begin{matrix} \} K \\ \} J-K \end{matrix} + \begin{bmatrix} \mathbf{n}_T \\ \cdots \\ 0 \end{bmatrix}. \qquad (14.3.13б)$$

Как видно из рис. 14.3.1, *б*, вектор $\mathbf{q}$ идентичен вектору наблюдаемого изображения $\mathbf{g}$ при учете только R его центральных элементов, причем $R = M - 2(L - 1)$. Остальные элементы $\mathbf{q}$ можно аппроксимировать следующим образом:

$$\mathbf{q} \approx \tilde{\mathbf{q}} = \mathbf{E}\mathbf{g}, \qquad (14.3.14)$$

где $\mathbf{E}$ — так называемая экстраполирующая матрица, которую мы определим ниже. Из формул (14.3.13) и (14.3.14) можно получить оценку $\mathbf{f}_T$ в виде

$$\hat{\mathbf{f}}_E = \mathbf{C}^{-1}\hat{\mathbf{q}}_E \qquad (14.3.15а)$$

или

$$\begin{matrix} K \{ \\ J-K \{ \end{matrix} \begin{bmatrix} \hat{\mathbf{f}}_T \\ \cdots \\ \cdot \\ \vdots \end{bmatrix} = [\mathbf{C}^{-1}] \begin{bmatrix} \tilde{\mathbf{q}} \\ \cdots \\ 0 \end{bmatrix}. \qquad (14.3.15б)$$

Алгебраическое сходство формул (14.3.2) и (14.3.15) позволяет непосредственно записать выражение для оценки как

$$\hat{\mathbf{f}}_E = [\mathbf{A}_J^{-1} \otimes \mathbf{A}_J^{-1}]\, \mathscr{C}^{-1} [\mathbf{A}_J \otimes \mathbf{A}_J]\, \tilde{\mathbf{q}}_E. \qquad (14.3.16)$$

Таким образом, вычислительный алгоритм для недоопределенной модели не будет отличаться от алгоритма (14.3.5)—(14.3.9) для переопределенной модели, если на этапе (б) матрицу наблюдаемого изображения $\mathbf{G}$ в выражении (14.3.6) заменить матрицей $\tilde{\mathbf{Q}} = (\mathbf{E} \otimes \mathbf{E})\, \mathbf{G}$.

Экстраполирующая матрица $\mathbf{E}$ предназначена для видоизменения элементов наблюдаемого изображения $\mathbf{g}$ таким образом, чтобы они аппроксимировали элементы $\mathbf{q}$. В рассматриваемом од-

номерном случае (см. рис. 14.3.1) справедливы следующие соотношения:

$$
\begin{aligned}
q(1) &= g(1) - h(3)f(1) - h(2)f(2),\\
q(2) &= g(2) - h(3)f(2),\\
q(m) &= g(m),\ 2 < m < M-2, \qquad (14.3.17)\\
q(M-1) &= g(M-1) - h(1)f(N-1),\\
q(M) &= g(M) - h(2)f(N-1) - h(1)f(N).
\end{aligned}
$$

Поскольку значения $\mathbf{f}$ неизвестны, систему уравнений (14.3.17) нельзя решить непосредственно. Однако можно принять допущение о непрерывности вектора исходных данных, согласно которому

$$f(1) = f(2) = f(3) \qquad (14.3.18а)$$

и

$$f(N-2) = f(N-1) = f(N), \qquad (14.3.18б)$$

и определить

$$q(1) = g(1)\,h(1)/S, \qquad (14.3.19а)$$

$$q(2) = g(2)\,\{[h(1) + h(2)]/S\}, \qquad (14.3.19б)$$

где $S = h(1) + h(2) + h(3)$. Аналогичные уравнения существуют для $q(M-1)$ и $q(M)$. В результате получаем экстраполирующую матрицу простого вида

$$\mathbf{E} = \left[\begin{array}{c:c:c} \alpha & 0 & 0 \\ \hdashline 0 & \mathbf{I} & 0 \\ \hdashline 0 & 0 & \beta \end{array}\right], \qquad (14.3.20)$$

где α и β — подматрицы размера $L \times L$, найденные с помощью уравнений (14.3.17) или их обобщений [20].

Рис. 14.3.2 иллюстрирует результат машинного моделирования алгоритма реставрации изображений методом псевдообращения матриц для недоопределенной модели в случае одномерной нерезкости. На первом этапе моделирования были отобраны K центральных элементов исходного изображения и на их основе сформирована совокупность векторов усеченного изображения $\mathbf{f}_T$ (рис. 14.3.2, *б*). Затем эти векторы были преобразованы с помощью гауссова импульсного отклика с $b_R = 1{,}5$ для моделирования нерезкости наблюдаемого изображения (рис. 14.3.2, *в*). Вид нерезкого изображения после обработки оператором экстраполяции $\mathbf{E}$ показан на рис. 14.3.2, *г*. На рис. 14.3.2, *д*, *е* представлены варианты исправленного изображения без обработки и с обработкой оператором экстраполяции. Полученные изображе-

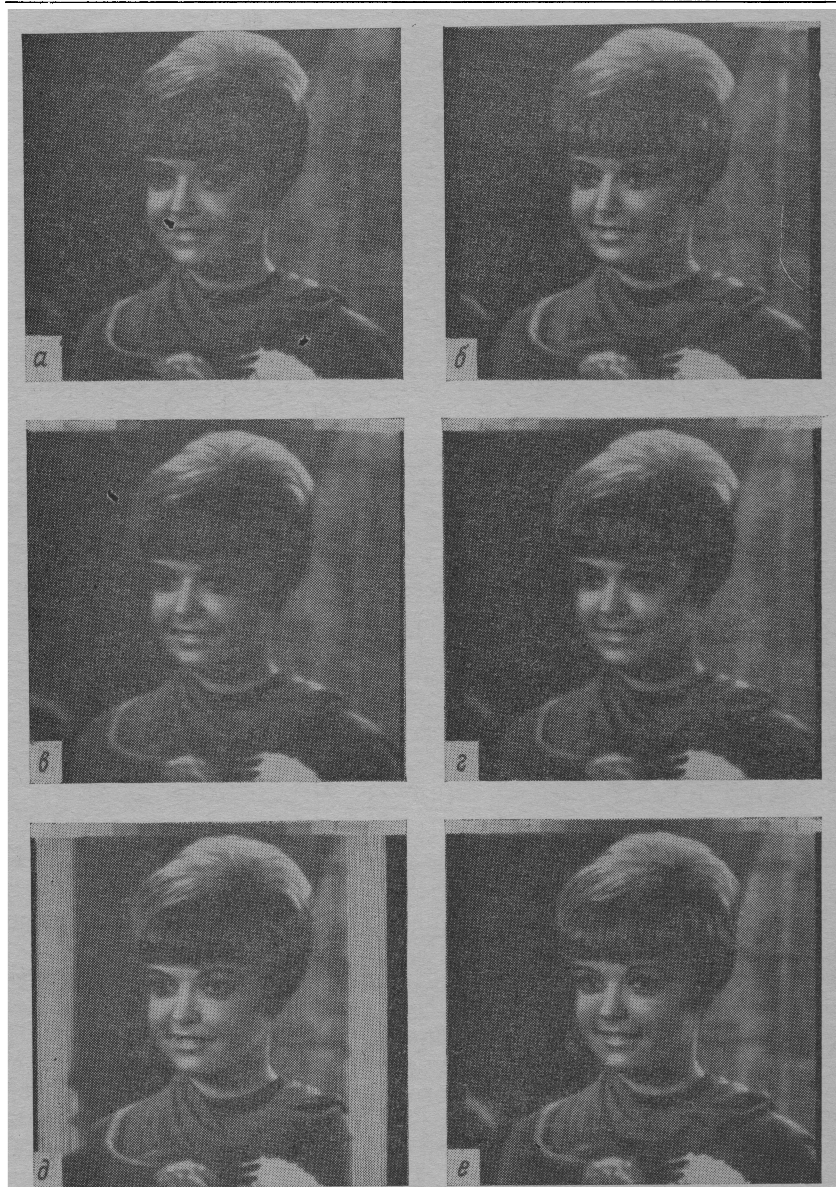


Рис. 14.3.2. Образцы изображений, исправленных методом псевдообращения матриц при незначительной гауссовой нерезкости ($b_R = 1{,}5$).

а — исходное изображение (вектор $\mathbf{f}$), *б* — усеченное изображение (вектор $\mathbf{f}_T$), *в* — наблюдаемое изображение (вектор $\mathbf{g}$), *г* — наблюдаемое изображение, обработанное оператором экстраполяции (вектор $\mathbf{q}$), *д* — исправленное изображение без экстраполяции (вектор $\hat{\mathbf{f}}_T$); *е* — исправленное изображение с экстраполяцией (вектор $\hat{\mathbf{f}}_T$).

ния наглядно демонстрируют важность экстраполяции. Если эта операция не производится, погрешности измерения на краях наблюдаемого изображения вызывают полное разрушение краев

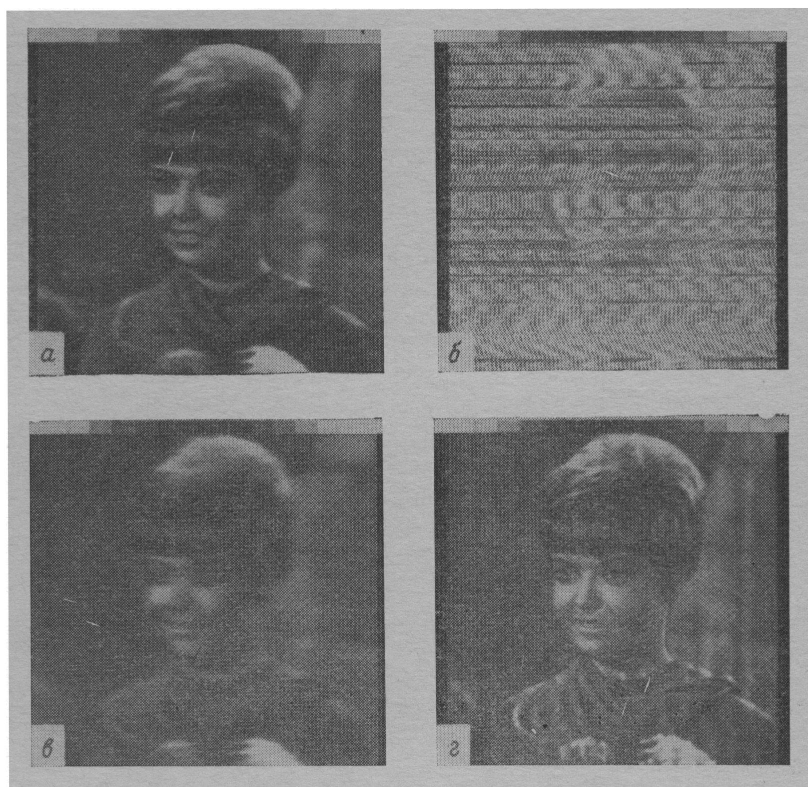


Рис. 14.3.3. Образцы изображений, исправленных методом псевдообращения матриц при средней и повышенной гауссовой нерезкости ($b_{\acute{R}} = 2$).

а — наблюдаемое изображение $\mathbf{g}$, *б* — исправленное изображение *а* ($\mathbf{f}_T$), *в* — наблюдаемое изображение ($\mathbf{g}$) с нерезкостью, вызванной равномерным движением ($L = 15$), *г* — исправленное изображение *в* ($\widehat{\mathbf{f}}_T$)

исправленного изображения. Напротив, реставрация с экстраполяцией обеспечивает вполне удовлетворительное субъективное качество изображения. Рис. 14.3.3 иллюстрирует результаты машинного эксперимента при повышенной нерезкости ($b_R = 2{,}0$), когда плохая обусловленность матрицы нерезкости становится сильно выраженной. При формировании видоизмененного наблюдаемого изображения посредством экстраполяции возникает большая остаточная ошибка, приводящая к непригодной

оценке (рис. 14.3.3, *б*). На рис. 14.3.3, *г* демонстрируется положительный эффект алгоритма реставрации методом псевдообращения матриц в случае горизонтальной нерезкости, вызванной движением (рис. 14.3.3, *в*). Поскольку в этом эксперименте импульсный отклик имеет прямоугольную форму, соответствующая матрица нерезкости обусловлена лучше, чем аналогичная матрица для гауссова импульсного отклика.

14.4. ПРОСТРАНСТВЕННАЯ РЕСТАВРАЦИЯ ИЗОБРАЖЕНИЙ МЕТОДОМ ПСЕВДООБРАЩЕНИЯ МАТРИЦ С ИСПОЛЬЗОВАНИЕМ СИНГУЛЯРНОГО РАЗЛОЖЕНИЯ

Как показано в гл. 5, любое дискретное изображение в матричной форме можно представить в виде ряда «собственных изображений», прибегнув к сингулярному разложению. Такой подход был применен для реставрации изображений [24—27]. В этом случае используется сингулярное разложение матрицы нерезкости, входящей в уравнение изображающей системы

$$\mathbf{g} = \mathbf{B}\mathbf{f} + \mathbf{n}. \tag{14.4.1}$$

Согласно формуле (5.2.3), матрицу нерезкости можно выразить в виде

$$\mathbf{B} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}^T, \tag{14.4.2}$$

где матрица $\mathbf{U}$ размера $P\times P$ и матрица $\mathbf{V}$ размера $Q\times Q$ — унитарные матрицы, составленные из собственных векторов матриц $\mathbf{B}\mathbf{B}^T$ и $\mathbf{B}^T\mathbf{B}$ соответственно, $\mathbf{\Lambda}$ — матрица размера $P\times Q$, диагональные элементы которой $\lambda(i)$ представляют собой собственные значения матриц $\mathbf{B}\mathbf{B}^T$ и $\mathbf{B}^T\mathbf{B}$. Ортогональность $\mathbf{U}$ и $\mathbf{V}$ позволяет представить матрицу нерезкости в виде ряда

$$\mathbf{B} = \sum_{i=1}^{R} \lambda^{1/2}(i)\,\mathbf{u}_i\mathbf{v}_i^T, \tag{14.4.3}$$

где $\mathbf{u}_i$ и $\mathbf{v}_i$ — i-е столбцы $\mathbf{U}$ и $\mathbf{V}$ соответственно, R — ранг матрицы $\mathbf{B}$.

Поскольку $\mathbf{U}$ и $\mathbf{V}$ являются унитарными матрицами, исходя из формулы (14.4.2) можно записать выражение для обобщенной обратной матрицы:

$$\mathbf{B}^{-} = \mathbf{V}\mathbf{\Lambda}^{-1/2}\mathbf{U}^T = \sum_{i=1}^{R} \lambda^{-1/2}(i)\,\mathbf{v}_i\mathbf{u}_i^T. \tag{14.4.4}$$

Рис. 14.4.1 иллюстрирует сингулярное разложение матрицы нерез кости. С помощью матрицы $\mathbf{B}^{-1}$ можно получить

$$\hat{\mathbf{f}} = \mathbf{B}^{-}\mathbf{g} = \mathbf{V}\mathbf{\Lambda}^{-1/2}\mathbf{U}^T\mathbf{g}, \tag{14.4.5a}$$

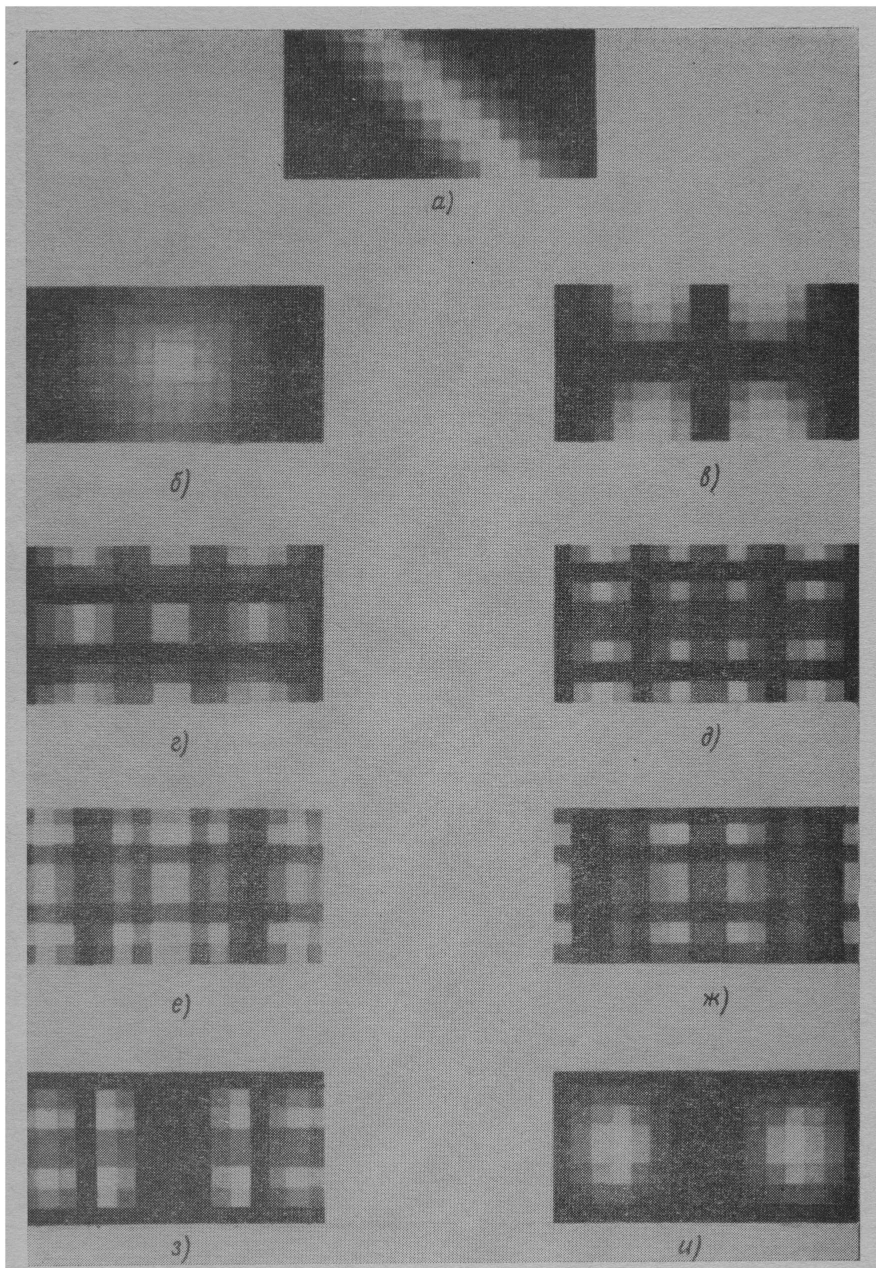


Рис 14.4.1. Изображения, соответствующие различным членам сингулярного разложения матрицы нерезкости ($b_R = 2{,}0$, $M = 8$, $N = 16$, $L = 9$).

а) $\mathbf{B}$, *б)* $\mathbf{u}_1\mathbf{v}_1^T$, $\lambda(1) = 0{,}871$, *в)* $\mathbf{u}_2\mathbf{v}_2^T$, $\lambda(2) = 0{,}573$; *г)* $\mathbf{u}_3\mathbf{v}_3^T$, $\lambda(3) = 0{,}285$; *д)* $\mathbf{u}_4\mathbf{v}_4^T$, $\lambda(4) = 0{,}108$, *е)* $\mathbf{u}_5\mathbf{v}_5^T$, $\lambda(5) = 0{,}034$; *ж)* $\mathbf{u}_6\mathbf{v}_6^T$, $\lambda(6) = 0{,}014$, *з)* $\mathbf{n}_7\mathbf{v}_7^T$, $\lambda(7) = 0{,}011$; *и)* $\mathbf{u}_8\mathbf{v}_8^T$, $\lambda(8) = 0{,}010$

или

$$\hat{\mathbf{f}} = \sum_{i=1}^{R} [\lambda(i)]^{-1/2} \mathbf{v}_i \mathbf{u}_i^T \mathbf{g} = \sum_{i=1}^{R} \lambda^{-1/2}(i) (\mathbf{u}_i^T \mathbf{g}) \mathbf{v}_i, \qquad (14.4.5б)$$

где учитывается тот факт, что скалярное произведение $\mathbf{u}_i^T \mathbf{g}$ дает в результате скаляр. Последнее уравнение лежит в основе последовательного оценивания: k-я оценка

$$\hat{\mathbf{f}}_k = \hat{\mathbf{f}}_{k-1} + [\lambda(k)]^{-1/2} (\mathbf{u}_k^T \mathbf{g}) \mathbf{v}_k. \qquad (14.4.6)$$

Одним из принципиальных достоинств последовательного подхода является то, что он позволяет решить проблему плохой обусловленности, которая обычно возникает только при высоком порядке сингулярных значений. При проведении реставрации в режиме диалога всегда можно прекратить разложение до появления ошибок, обусловленных неустойчивостью вычислительного процесса.

На рис. 14.4.2 представлены образцы изображений, получаемых на различных этапах реставрации с использованием сингулярного разложения испытательного изображения, представленного на рис. 13.7.3 (недоопределенная модель с плохо обусловленной матрицей гауссовой нерезкости). Результатом одношагового псевдообращения матрицы является изображение, полностью разрушенное ошибками вычислений. На шестом этапе обеспечивается наилучшее субъективное качество исправленного изображения, которое заметно лучше нерезкого исходного изображения. Тем не менее самая малая ошибка по критерию наименьших квадратов достигается при трех сингулярных значениях.

Возможности метода реставрации изображений на основе сингулярного разложения в соответствии с формулами (14.4.5) и (14.4.6) ограничены главным образом сложностью вычислений. Сначала определяются собственные векторы $\mathbf{u}_i$ и $\mathbf{v}_i$ матриц $\mathbf{B}\mathbf{B}^T$ и $\mathbf{B}^T\mathbf{B}$. Затем проводятся векторные вычисления согласно формулам (14.4.5) и (14.4.6). Даже если матрицу $\mathbf{B}$ можно представить в виде прямого произведения, что позволяет производить раздельное псевдообращение строк и столбцов, вычисления в общем случае оказываются чрезвычайно трудоемкими.

Вычислительный алгоритм на основе псевдообращения матриц, описанный в предыдущем разделе, удается приспособить для решения задачи реставрации изображений на основе сингулярного разложения в особом случае пространственно-инвариантной нерезкости [21]. Пользуясь моделью, определяемой формулой (14.3.13)

$$\mathbf{q}_E = \mathbf{C}\mathbf{f}_E + \mathbf{n}_E, \qquad (14.4.7)$$

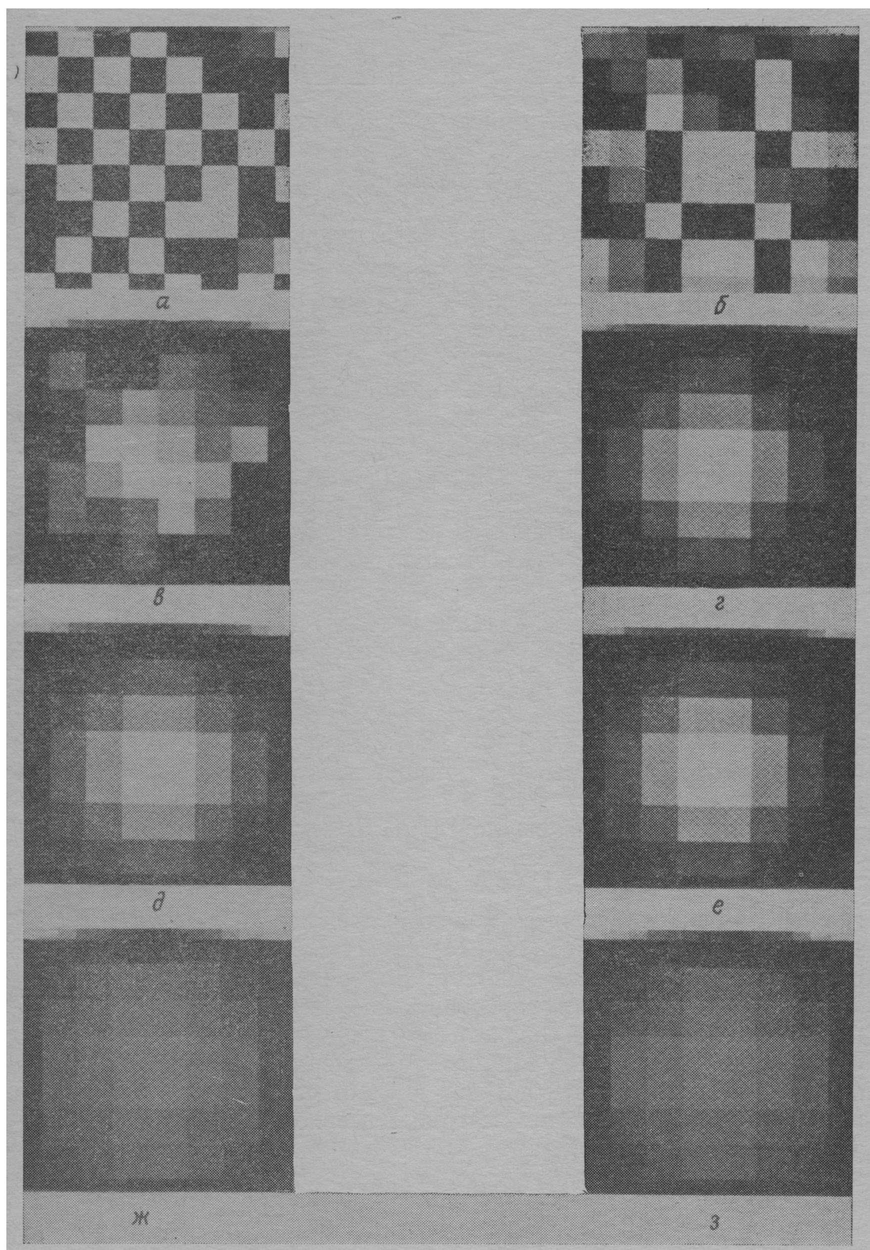


Рис. 14.4.2. Образцы испытательных изображений, исправленных методом сингулярного разложения. Нерезкость внесена с помощью импульсного отклика гауссовой формы ($b_R = b_C = 1{,}2$). Случай недоопределенной модели ($M = 8$, $N = 12$, $L = 5$). Наблюдаемое изображение с шумом (дисперсия равна 10 уровням шкалы).

а — 8 сингулярных значений (СЗ), МСКО = 2695,81 %; *б* — 7 СЗ, МСКО = 148,93 %; *в* — 6 СЗ, МСКО = 6,88 %; *г* — 5 СЗ, МСКО = 3,31 %; *д* — 4 СЗ, МСКО = 3,06 %; *е* — 3 СЗ, МСКО = 3,05 %; *ж* — 2 СЗ, МСКО = 9,52 %; *з* — 1 СЗ, МСКО = 9,52 %.

можно получить сингулярное разложение циклической матрицы $\mathbf{C}$:

$$\mathbf{C} = \mathbf{X}\mathbf{\Delta}^{1/2}\mathbf{Y}^{*T}, \tag{14.4.8}$$

где $\mathbf{X}$ и $\mathbf{Y}$ — унитарные матрицы, определяемые соотношениями

$$\mathbf{X}\,[\mathbf{C}\mathbf{C}^T]\,\mathbf{X}^{*T} = \mathbf{\Delta} \tag{14.4.9а}$$

и

$$\mathbf{Y}\,[\mathbf{C}^T\mathbf{C}]\,\mathbf{Y}^{*T} = \mathbf{\Delta}. \tag{14.4.9б}$$

Поскольку $\mathbf{C}$ — циклическая матрица, $\mathbf{C}\mathbf{C}^T$ также является циклической матрицей. Следовательно, $\mathbf{X}$ и $\mathbf{Y}$ должны быть эквивалентны матрице преобразования Фурье $\mathbf{A}$ (или $\mathbf{A}^{-1}$), так как эта матрица производит диагонализацию циклической матрицы. С целью стандартизации примем $\mathbf{X} = \mathbf{Y} = \mathbf{A}^{-1}$. В результате получим собственные векторы $\mathbf{x}_i = \mathbf{y}_i$ — строки матриц $\mathbf{X}$ и $\mathbf{Y}$ — в виде комплексных экспоненциальных базисных функций

$$x_k^*(j) = \exp\left\{\frac{2\pi i}{J}(k-1)(j-1)\right\} \tag{14.4.10}$$

преобразования Фурье при $1 \leqslant j,\ k \leqslant J$. Более того, находим

$$\mathbf{\Delta} = \mathscr{C}\mathscr{C}^{*T}, \tag{14.4.11}$$

где $\mathscr{C}$ — матрица двумерной циклической свертки в спектральной области.

В соответствии с формулой (14.4.5) можно записать

$$\hat{\mathbf{f}}_E = \mathbf{A}^{-1}\mathbf{\Delta}^{-1/2}\mathbf{A}\tilde{\mathbf{q}}_E, \tag{14.4.12}$$

где $\tilde{\mathbf{q}}_E$ — видоизмененное наблюдаемое нерезкое изображение, заданное выражениями (14.3.14) и (14.3.15). Можно заметить, что формула (14.4.12) является точным эквивалентом формулы (14.3.16), которая определяет оценку на основе псевдообращения в спектральной области. Последовательная реставрация с использованием сингулярного разложения, аналогичная процедуре в соответствии с формулой (14.4.6), описывается уравнением (14.4.12) при замене в нем матрицы псевдообращения $\mathbf{\Delta}^{-1/2}$ оператором

$$\mathbf{\Delta}_T^{-1/2} = \begin{bmatrix} [\Delta_T(1)]^{-1/2} & & & & \\ & [\Delta_T(2)]^{-1/2} & & 0 & \\ & & [\Delta_T(T)]^{-1/2} & & \\ & 0 & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}. \tag{14.4.13}$$

Не во всех случаях необходимо полностью отбрасывать высокочастотные члены, чтобы избежать ошибок вследствие плохой обусловленности. Усечение можно заменить альтернативным мето-



Рис. 14.4 3. Образцы изображений, исправленных методом псевдообращения матриц с использованием сингулярного разложения при одномерной (горизонтальной) гауссовой нерезкости ($b_R = 3$, $L = 23$, $J = 256$).

а — наблюдаемое нерезкое изображение, *б* — исправленное изображение при $T = 58$; *в* — исправленное изображение *а* при $T = 60$.

дом, при котором диагональные нулевые элементы заменяются элементами в виде $[\Delta\,(T)]^{-1/2}$ или даже некоторой последовательностью, убывающей по мере повышения частоты. По существу этот прием аналогичен методу усеченной инверсной фильтрации согласно формуле (14.1.11) для непрерывных изображений.

На рис. 14.4.3 приведены образцы изображений, исправленных методом псевдообращения матриц на основе сингулярного разложения при одномерной гауссовой нерезкости ($b_R = 3{,}0$). Следует отметить, что попытка реставрировать нерезкое изображение ме-

тодом обычного псевдообращения приводит к недопустимо большим ошибкам из-за плохой обусловленности даже при коэффициенте нерезкости $b_R = 2{,}0$ (рис. 14.3.3, *б*).

14.5. ПРОСТРАНСТВЕННАЯ РЕСТАВРАЦИЯ ИЗОБРАЖЕНИЙ МЕТОДОМ РЕГРЕССИИ

Методы реставрации изображений на основе псевдообращения матриц имеют одно принципиальное ограничение: шум наблюдаемого изображения может стать причиной резко выраженной неустойчивости вычислительного процесса и сделать непригодной полученную оценку изображения. При решении ряда задач эту проблему можно обойти, прибегнув к методам реставрации на основе регрессии, использующим определенную априорную информацию о статистических свойствах наблюдаемого шума [19].

Рассмотрим матричное уравнение

$$\mathbf{g} = \mathbf{B}\mathbf{f} + \mathbf{n}, \tag{14.5.1}$$

описывающее нерезкое изображение с аддитивным шумом, где $\mathbf{B}$ — матрица нерезкости размера $P \times Q$; предполагается, что шум имеет нулевое среднее значение и известную ковариационную матрицу $\mathbf{K}_n$. Метод реставрации на основе регрессии предусматривает нахождение оценки

$$\hat{\mathbf{f}} = \mathbf{W}\mathbf{g}, \tag{14.5.2}$$

где $\mathbf{W}$ — реставрирующая матрица, минимизирующая взвешенную ошибку

$$0(\hat{\mathbf{f}}) = (\mathbf{g} - \mathbf{B}\hat{\mathbf{f}})^T \mathbf{K}_n^{-1} (\mathbf{g} - \mathbf{B}\hat{\mathbf{f}}). \tag{14.5.3}$$

Ошибку реставрации можно минимизировать классическим методом приравнивания нулю частной производной $\theta(\hat{\mathbf{f}})$ по $\hat{\mathbf{f}}$. Таким образом,

$$\partial\theta(\hat{\mathbf{f}})/\partial\hat{\mathbf{f}} = 0 = -2\mathbf{B}^T \mathbf{K}_n^{-1} (\mathbf{g} - \mathbf{B}\hat{\mathbf{f}}). \tag{14.5.4}$$

Если $P > Q$, а обратная матрица относительно $\mathbf{B}^T \mathbf{K}_n^{-1} \mathbf{B}$ существует, результирующая реставрирующая матрица принимает вид

$$\mathbf{W} = (\mathbf{B}^T \mathbf{K}_n^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{K}_n^{-1}. \tag{14.5.5}$$

Полученный оператор реставрации применим к переопределенной модели наблюдаемого изображения. Если шум белый с дисперсией σ_n^2, $\mathbf{K}_n = \sigma_n^2 \mathbf{I}$ и матрица регрессии становится обобщенной обратной матрицей (14.2.9а) ранга Q, соответствующей переопределенной системе.

Можно показать [25], что в случае недоопределенной модели ($P < Q$) оценка на основе регрессии имеет вид

$$\hat{\mathbf{f}} = (\mathbf{K}^{-1}\mathbf{B})^{-} \mathbf{K}^{-1}\mathbf{g} + [\mathbf{I} - (\mathbf{K}^{-1}\mathbf{B})^{-} \mathbf{K}^{-1}\mathbf{B}]\, \mathbf{v}, \qquad (14.5.6)$$

где $\mathbf{v}$ — произвольный вектор, $\mathbf{K}$ — матрица, определяемая соотношением

$$\mathbf{K}_n = \mathbf{K}\mathbf{K}^T, \qquad (14.5.7)$$

$\mathbf{K}_n$ — ковариационная матрица шума. Таким образом, в недоопределенной модели реставрации оценка не является единственной. Оценка с наименьшей нормой есть просто

$$\hat{\mathbf{f}} = (\mathbf{K}^{-1}\mathbf{B})^{-} \mathbf{K}^{-1}\mathbf{g}, \qquad (14.5.8)$$

поэтому результирующий оператор реставрации на основе регрессии сводится к

$$\mathbf{W} = (\mathbf{K}^{-1}\mathbf{B})^{-} \mathbf{K}^{-1}. \qquad (14.5.9)$$

Если шум белый, $\mathbf{K} = \sigma_n^2\mathbf{I}$ и этот оператор принимает вид обобщенной обратной матрицы (14.2.9б) ранга P для недоопределенной модели.

14.6. ПРОСТРАНСТВЕННАЯ РЕСТАВРАЦИЯ ИЗОБРАЖЕНИЙ МЕТОДОМ ВИНЕРОВСКОГО ОЦЕНИВАНИЯ

При использовании метода пространственной реставрации изображений на основе регрессии шумовое поле моделируется некоторой реализацией двумерного случайного процесса с известными средним и ковариационной функцией. В методах винеровской оценки, кроме того, предполагается, что идеальное изображение также является реализацией двумерного случайного процесса с известными первым и вторым моментами [19, 20, 29].

14.6.1. ВИНЕРОВСКОЕ ОЦЕНИВАНИЕ В ОБЩЕМ СЛУЧАЕ

Рассмотрим обобщенную дискретную модель, представленную на рис. 14.6.1, в которой вектор изображения $\mathbf{f}$ размера $Q\times 1$ подвергается каким-либо искажениям (поэлементным и пространственным), в результате чего образуется вектор размера $P\times 1$ наблюдаемого изображения $\mathbf{g}$. Формируемая оценка $\mathbf{f}$ описывается линейной операцией вида

$$\hat{\mathbf{f}} = \mathbf{W}\mathbf{g} + \mathbf{b}, \qquad (14.6.1)$$

где $\mathbf{W}$ — реставрирующая матрица размера $Q\times P$, $\mathbf{b}$ — вектор смещения оценки размера $Q\times 1$. Цель винеровского оценивания

сводится к выбору **W** и **b**, минимизирующих среднеквадратическую ошибку реставрации, которую можно определить как

$$\mathscr{E} = E\{(\mathbf{f} - \hat{\mathbf{f}})^T (\mathbf{f} - \hat{\mathbf{f}})\} \tag{14.6.2а}$$

или

$$\mathscr{E} = \mathrm{tr}\,[E\{(\mathbf{f} - \hat{\mathbf{f}})(\mathbf{f} - \hat{\mathbf{f}})^T\}]. \tag{14.6.2б}$$

В формуле (14.6.2а) ошибка определена через скалярное произведение и равна сумме квадратов элементов вектора ошибки $(\mathbf{f}-\hat{\mathbf{f}})$;

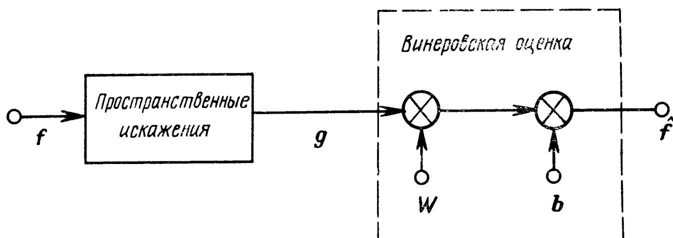


Рис. 14.6.1. Пространственная реставрация изображений методом винеровского оценивания.

формула (14.6.2б) предусматривает составление ковариационной матрицы вектора ошибки и последующее суммирование ее дисперсий (диагональных элементов) при нахождении следа. Независимо от способа представления ошибки (14.6.2) ее можно минимизировать, дифференцируя $\mathscr{E}$ по $\hat{\mathbf{f}}$. Другой подход, обладающий универсальной полезностью, основан на использовании принципа ортогональности [30, стр. 219]: с его помощью находят значения **W** и **b**, минимизирующие среднеквадратическую ошибку. В теории реставрации изображений принцип ортогональности позволяет сформулировать следующие два необходимых и достаточных условия минимизации среднеквадратической ошибки реставрации:

1. Математическое ожидание оценки изображения должно быть равно математическому ожиданию исходного изображения, т. е.

$$E\{\hat{\mathbf{f}}\} = E\{\mathbf{f}\}. \tag{14.6.3}$$

2. Ошибка реставрации должна быть ортогональна к «центрированному» наблюдаемому изображению относительно его среднего, т. е.

$$E\{(\hat{\mathbf{f}} - \mathbf{f})(\mathbf{g} - E\{\mathbf{g}\})^T\} = \mathbf{0}. \tag{14.6.4}$$

Из условий 1 и 2 получаем

$$\mathbf{b} = E\{\mathbf{f}\} - \mathbf{W}E\{\mathbf{g}\} \tag{14.6.5}$$

и

$$E\{(\mathbf{W}\mathbf{g} + \mathbf{b} - \mathbf{f})(\mathbf{g} - E\{\mathbf{g}\})^T\} = \mathbf{0} \tag{14.6.6}$$

соответственно. Подстановка в формулу (14.6.6) выражения (14.6.5) для вектора смещения ошибки $\mathbf{b}$ дает после ряда упрощений

$$\mathbf{W} = \mathbf{K}_{fg}(\mathbf{K}_{gg})^{-1}, \tag{14.6.7}$$

где $\mathbf{K}_{gg}$ — ковариационная матрица размера $P \times P$ вектора наблюдаемого изображения (которая предполагается невырожденной), $\mathbf{K}_{fg}$ — взаимно ковариационная матрица размера $Q \times P$ для исходного и наблюдаемого изображений. Таким образом, оптимальный вектор смещения $\mathbf{b}$ и реставрирующую матрицу $\mathbf{W}$ можно определить непосредственно по смешанным моментам первого и второго порядка векторов идеального и наблюдаемого изображений. Следует отметить, что найденное решение справедливо для нелинейных и пространственно-зависимых искажений. Следующие разделы посвящены отдельным частным случаям применения винеровского оценивания.

14.6.2. ВИНЕРОВСКОЕ ОЦЕНИВАНИЕ В СЛУЧАЕ НЕРЕЗКОГО ИЗОБРАЖЕНИЯ С ШУМОМ

В дискретной модели, описывающей нерезкое изображение с аддитивным шумом, используется вектор наблюдаемого изображения вида

$$\mathbf{g} = \mathbf{B}\mathbf{f} + \mathbf{n}. \tag{14.6.8}$$

В такой модели винеровская оценка составляется из вектора смещения

$$\mathbf{b} = E\{\mathbf{f}\} - \mathbf{W}E\{\mathbf{g}\} = E\{\mathbf{f}\} - \mathbf{W}\mathbf{B}E\{\mathbf{f}\} + \mathbf{W}E\{\mathbf{n}\} \tag{14.6.9}$$

и матричного оператора

$$\mathbf{W} = \mathbf{K}_{fg}(\mathbf{K}_{gg})^{-1} = \mathbf{K}_f\mathbf{B}^T(\mathbf{B}\mathbf{K}_f\mathbf{B}^T + \mathbf{K}_n)^{-1} \tag{14.6.10}$$

или на основе матричного тождества (5.1.11)

$$\mathbf{W} = (\mathbf{B}^T\mathbf{K}_n^{-1}\mathbf{B} + \mathbf{K}_f^{-1})^{-1}\mathbf{B}^T\mathbf{K}_n^{-1}. \tag{14.6.11}$$

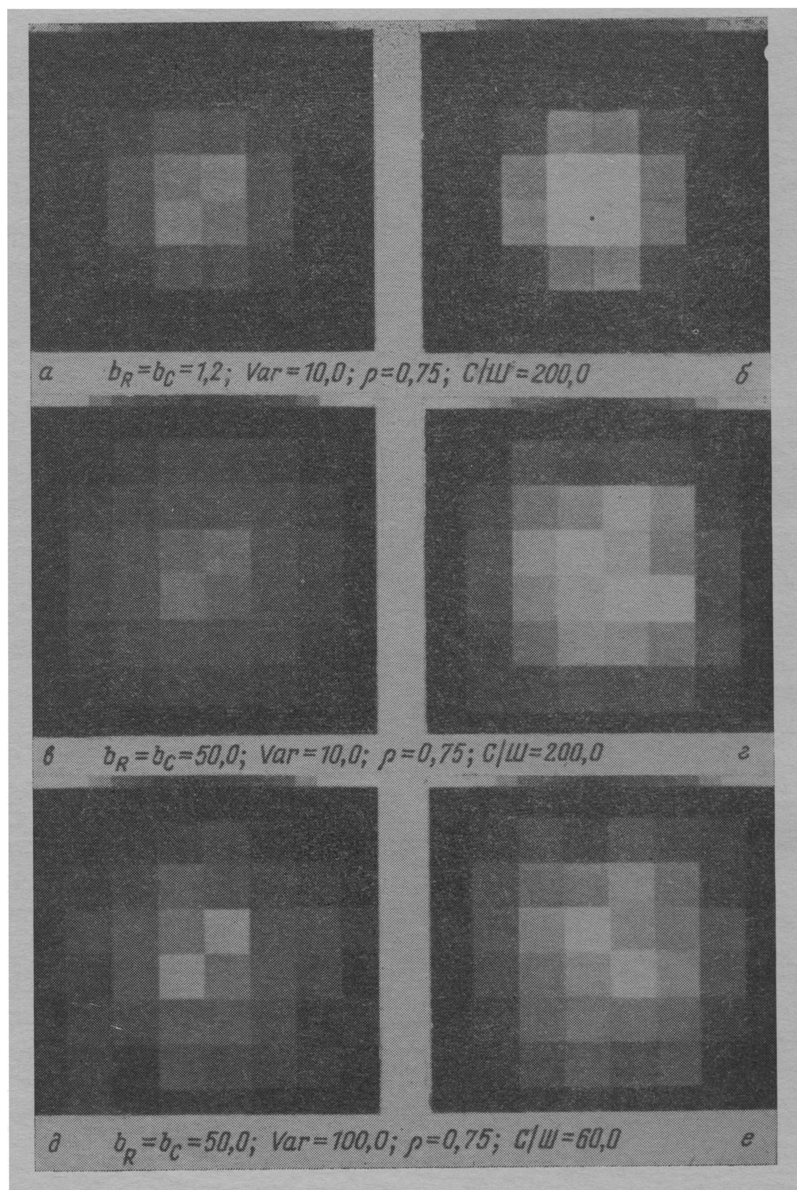


Рис. 14.6.2. Образцы испытательных изображений, исправленных методом винеровского оценивания. Нерезкость внесена с помощью импульсного отклика гауссовой формы. Случай недоопределенной модели ($M = 8$, $N = 12$, $L = 5$).

а — нерезкое изображение (МСКО = 4,91 %), *б* — исправленное изображение *а* (МСКО = = 3,17%); *в* — нерезкое изображение (МСКО = 7,99 %), *г* — исправленное изображение *в* (МСКО = 4,20 %); *д* — нерезкое изображение (МСКО = 7,93 %), *е* — исправленное изображение *д* (МСКО = 4,74 %).

Уравнения (14.6.10) и (14.6.11) эквивалентны, однако первым из них следует пользоваться для недоопределенных систем, вторым — для переопределенных систем. При этом обеспечивается наименьший размер обратных матриц.

Если предполагается, что идеальное изображение не обладает корреляцией, то $\mathbf{K}_f = \sigma_f^2\mathbf{I}$, где σ_f^2 представляет энергию изображения. Тогда выражение (14.6.10) принимает вид

$$\mathbf{W} = \sigma_f^2\mathbf{B}^T(\sigma_f^2\mathbf{B}\mathbf{B}^T + \mathbf{K}_n)^{-1}. \qquad (14.6.12)$$

При том же условии формула (14.6.11) переходит в

$$\mathbf{W} = (\mathbf{B}^T\mathbf{K}_n^{-1}\mathbf{B} + \sigma_f^{-2}\mathbf{I})^{-1}\mathbf{B}^T\mathbf{K}_n^{-1}. \qquad (14.6.13)$$

Для случая белого шума с энергией σ_n^2 получаем

$$\mathbf{W} = \mathbf{B}^T[\mathbf{B}\mathbf{B}^T + (\sigma_n^2/\sigma_f^2)\mathbf{I}]^{-1} \qquad (14.6.14)$$

и

$$\mathbf{W} = [\mathbf{B}^T\mathbf{B} + (\sigma_f^{-2}/\sigma_n^{-2})\mathbf{I}]^{-1}\mathbf{B}^T. \qquad (14.6.15)$$

Если отношение энергии изображения к энергии шума $(\sigma_f/\sigma_n)^2$ стремится к бесконечности, то винеровская оценка (14.6.14) переходит в оценку обобщенного обращения для недоопределенной модели, а оценка (14.6.15) — в оценку обобщенного обращения для переопределенной модели.

На рис. 14.6.2 показаны образцы изображений, полученных при реставрации изображения рис. 13.7.3 с гауссовой нерезкостью (случай недоопределенной модели). Примеры винеровской реставрации изображений большого размера с использованием вычислительного алгоритма Прэтта и Даварианa [20] представлены на рис. 14.6.3. На рис. 14.6.3, *а* нерезкость вызвана движением по горизонтали, причем был использован прямоугольный импульсный отклик длиной $L = 11$. Исправленная центральная область, содержащая 129 элементов, вставлена в нерезкое изображение, чтобы продемонстрировать повышение субъективного качества изображения, обеспечиваемое реставрацией. Изображение, показанное на рис. 14.6.3, *в*, сформировано электронным микроскопом; уровень шума и импульсный отклик нерезкости были определены непосредственно по фотографическому изображению

методами, описанными в гл. 15. Эти данные позволили реставрировать центральную область размером 129×129 элементов, которая затем была вставлена в исходное нерезкое изображение.

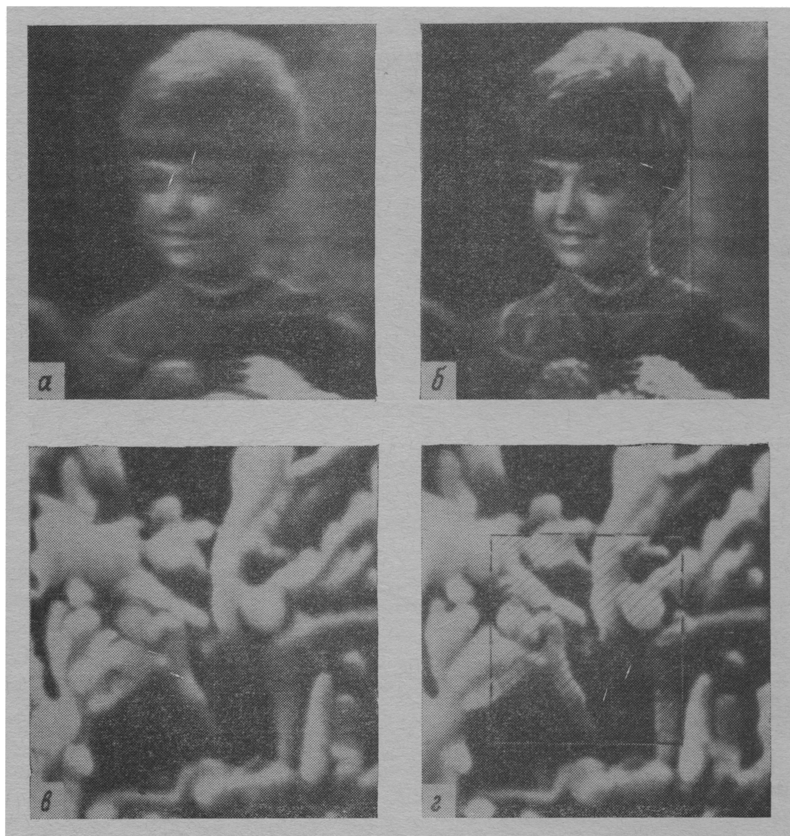


Рис. 14.6.3. Образцы изображений, исправленных методом винеровского оценивания.

а — наблюдаемое нерезкое изображение при горизонтальном движении ($L = 11$), *б* — исправленное изображение *а*; *в* — наблюдаемое изображение, формируемое электронным микроскопом; *г* — исправленное изображение *в*

14.7. МЕТОДЫ РЕСТАВРАЦИИ ИЗОБРАЖЕНИЙ СО СГЛАЖИВАНИЕМ

Проблемы плохой обусловленности, возникающие при реставрации изображений, иногда удается преодолеть, прибегая к методам сглаживания и регуляризации [31—33], которые в основном связаны с введением меры гладкости решений, полученных методом наименьших квадратов.

Методы сглаживания имеют две формулировки [19]. Согласно первой из них, ищется минимум величины $\hat{\mathbf{f}}^T\mathbf{S}\hat{\mathbf{f}}$ при ограничении

$$(\mathbf{g} - \mathbf{B}\hat{\mathbf{f}})^T \mathbf{M} (\mathbf{g} - \mathbf{B}\hat{\mathbf{f}}) = e, \tag{14.7.1}$$

где $\mathbf{S}$ — сглаживающая матрица, $\mathbf{M}$ — матрица взвешивания ошибки, e — остаточная скалярная ошибка оценки. Матрица $\mathbf{M}$ часто выбирается равной обратной матрице относительно ковариационной матрицы шума, т. е. $\mathbf{M} = \mathbf{K}_n^{-1}$. Обычной формой сглаживающей матрицы является матрица

$$\mathbf{S} = \mathbf{S}_1 \otimes \mathbf{S}_1, \tag{14.7.2}$$

где

$$\mathbf{S}_1 = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 & & & & & 0 \\ -2 & 5 & -4 & 1 & 0 & 0 & 0 & & & & & \cdot \\ 1 & -4 & 6 & -4 & 1 & 0 & 0 & & & & & \cdot \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & & & & & \cdot \\ 0 & 0 & 1 & -4 & 6 & -4 & 1 & & & & & \cdot \\ & & & & & \ddots & & & & & & \\ & & & & & 1 & -4 & 6 & -4 & 1 & & 0 \\ & & & & & 0 & 1 & -4 & 6 & -4 & & 1 \\ & & & & & 0 & 0 & 1 & -4 & 5 & & -2 \\ 0 & & & & & 0 & 0 & 0 & 1 & -2 & & 1 \end{bmatrix} \tag{14.7.3}$$

есть одномерный линейный оператор, формирующий скользящее окно для получения оценки. Условный минимум гладкости совпадает с минимумом функции Лагранжа

$$F(\hat{\mathbf{f}}, \lambda) = \hat{\mathbf{f}}^T\mathbf{S}\mathbf{f} + \lambda\,[(\mathbf{g} - \mathbf{B}\hat{\mathbf{f}})\,\mathbf{M}^T(\mathbf{g} - \mathbf{B}\hat{\mathbf{f}}) - e]. \tag{14.7.4}$$

Вычисляя производные по $\mathbf{f}$ и λ и приравнивая их нулю, находим оценки

$$\hat{\mathbf{f}} = [\mathbf{B}\mathbf{M}^T\mathbf{B} + (1/\lambda)\mathbf{S}]^{-1}\,\mathbf{B}^T\mathbf{M}\mathbf{g} \tag{14.7.5а}$$

и

$$\hat{\mathbf{f}} = \mathbf{S}^{-1}\mathbf{B}^T\,[\mathbf{B}\mathbf{S}^{-1}\mathbf{B}^T + (1/\lambda)\,\mathbf{M}^{-1}]^{-1}\,\mathbf{g} \tag{14.7.5б}$$

для несингулярной переопределенной и несингулярной недоопределенной систем соответственно. В выражениях (14.7.5) лагранжев коэффициент λ должен удовлетворять условию (14.7.1), т. е. предполагается, что достигается компромисс между остаточной ошибкой и гладкостью оценки.

Теперь рассмотрим вторую формулировку. В этом случае решается задача минимизации среднеквадратической ошибки $(\mathbf{g} - \mathbf{B}\hat{\mathbf{f}})^T \mathbf{M} (\mathbf{g} - \mathbf{B}\hat{\mathbf{f}})$ при ограничении

$$\hat{\mathbf{f}}^T \mathbf{S} \hat{\mathbf{f}} = d, \tag{14.7.6}$$

где d — скаляр, соответствующий фиксированной гладкости. Функция Лагранжа принимает вид

$$G(\hat{\mathbf{f}}, \gamma) = (\mathbf{g} - \mathbf{B}\hat{\mathbf{f}}) \mathbf{M}^T (\mathbf{g} - \mathbf{B}\hat{\mathbf{f}}) + \gamma (\hat{\mathbf{f}}^T \mathbf{S} \hat{\mathbf{f}} - d). \tag{14.7.7}$$

По этому уравнению находится оптимальное решение для переопределенной несингулярной системы:

$$\hat{\mathbf{f}} = (\mathbf{B}^T \mathbf{M} \mathbf{B} + \gamma \mathbf{S})^{-1} \mathbf{B}^T \mathbf{M} \mathbf{g}. \tag{14.7.8а}$$

Оценка для недоопределенной системы принимает вид

$$\hat{\mathbf{f}} = \mathbf{S}^{-1} \mathbf{B}^T (\mathbf{B} \mathbf{S}^{-1} \mathbf{B}^T + \gamma \mathbf{M}^{-1})^{-1} \mathbf{g}. \tag{14.7.8б}$$

Сравнение формул (14.7.5) и (14.7.8) показывает, что для обеих взаимно обратных задач получаются решения, которые различаются только взаимно обратными множителями Лагранжа.

Оценки (14.7.8), полученные методом сглаживания, имеют близкое сходство с оценками на основе регрессии и винеровскими оценками, найденными выше. При $\gamma = 0$, $\mathbf{S} = \mathbf{I}$ и $\mathbf{M} = \mathbf{K}_n^{-1}$, где $\mathbf{K}_n$ — ковариационная матрица шума, оценки, полученные методами сглаживания и регрессии, становятся эквивалентными. При $\gamma = 1$, $\mathbf{S} = \mathbf{K}_f^{-1}$ и $\mathbf{M} = \mathbf{K}_n^{-1}$, где $\mathbf{K}_f$ — ковариационная матрица изображения, оценка, полученная методом сглаживания, эквивалентна винеровской оценке. Это сходство позволяет объяснить, почему реставрация методом регрессии и винеровская реставрация обеспечивают более гладкие оценки, чем реставрация методом псевдообращения матриц. При использовании методов сглаживания и регуляризации смещение решения удается определить только в виде функциональной зависимости от $\mathbf{f}$, даже если можно вычислить дисперсию решения.

Как уже указывалось, Хант [18] разработал фильтр для реставрации изображений в спектральной области с использованием критерия наименьших квадратов, имеющий частотную характеристику (14.1.22). Это выражение можно использовать как основу быстрого вычислительного алгоритма для осуществления процедуры реставрации с использованием сглаживания согласно (14.7.8), если соответствующим образом доопределить наблюдаемый вектор $\mathbf{g}$.

14.8. МЕТОДЫ РЕСТАВРАЦИИ С ИСПОЛЬЗОВАНИЕМ ОГРАНИЧЕНИЙ

Для повышения качества исправленных изображений в моделях с плохой обусловленностью было предложено [19] вводить ограничения в виде равенств и ограничения в виде неравенств. Можно, например, задавать значения отдельных элементов, отношения отдельных элементов, сумму части или всех элементов или же предельно допустимые уровни элементов.

В переопределенной линейной модели могут использоваться ограничения в виде системы линейных уравнений

$$\mathbf{A}\hat{\mathbf{f}} = \mathbf{t}, \tag{14.8.1}$$

где $\mathbf{A}$ — заданная матрица размера $P \times N$ ранга $P < N$, описывающая ограничения, $\hat{\mathbf{f}}$ — оценка при наличии ограничений, $\mathbf{t}$ — известный вектор размера $P \times 1$. Минимизация взвешенной ошибки $(\mathbf{g} - \mathbf{Bf})\ \mathbf{K}_n^{-1}\ (\mathbf{g} - \mathbf{Bf})$ с учетом ограничений (14.8.1) позволяет получить следующую оценку [28, стр. 100]:

$$\hat{\mathbf{f}} = \tilde{\mathbf{f}} + (\mathbf{B}^T \mathbf{K}_n^{-1} \mathbf{B})^{-1} \mathbf{A}^T [\mathbf{A} (\mathbf{B}^T \mathbf{K}_n^{-1} \mathbf{B})^{-1} \mathbf{A}^T]^{-1} (\mathbf{t} - \mathbf{A}\tilde{\mathbf{f}}), \tag{14.8.2}$$

где $\tilde{\mathbf{f}}$ — решение без ограничений. Ковариационная матрица оптимального решения определяется как

$$\mathbf{K}_{\hat{f}} = (\mathbf{B}^T \mathbf{K}_n^{-1} \mathbf{B})^{-1} - (\mathbf{B}^T \mathbf{K}_n^{-1} \mathbf{B})^{-1} \mathbf{A}^T [\mathbf{A} (\mathbf{B}^T \mathbf{K}_n^{-1} \mathbf{B}) \mathbf{A}^T]^{-1} \times$$
$$\times \mathbf{A} (\mathbf{B}^T \mathbf{K}_n^{-1} \mathbf{B})^{-1}, \tag{14.8.3}$$

где $\mathbf{K}_{\tilde{f}} = (\mathbf{B}^T \mathbf{K}_n^{-1} \mathbf{B})$ — (положительно определенная) ковариационная матрица оценки без ограничений. Можно показать, что вторая матрица в выражении (14.8.3) является неотрицательно определенной матрицей ранга P. Следовательно, каждый диагональный элемент матрицы $\mathbf{K}_{\hat{f}}$ меньше или равен соответствующему элементу $\mathbf{K}_{\tilde{f}}$, а дисперсия каждого элемента вектора оценки с ограничениями меньше дисперсии соответствующего элемента вектора оценки без ограничений. Однако это не означает, что оценка с ограничениями всегда лучше оценки без ограничений. Более того, первая может оказаться смещенной, вторая — нет. Смещение решения при наличии ограничений

$$\mathbf{b} = \mathbf{f} - E\{\hat{\mathbf{f}}\} = -(\mathbf{b}^T \mathbf{K}_n^{-1} \mathbf{B})^{-1} \mathbf{A}^T [\mathbf{A} (\mathbf{B}^T \mathbf{K}_n^{-1} \mathbf{B})^{-1} \mathbf{A}^T]^{-1} [\mathbf{t} - \mathbf{Af}] \tag{14.8.4}$$

будет иметь нулевое значение тогда и только тогда, когда ограничения удовлетворяются истинным вектором $\mathbf{f}$, т. е. тогда и только тогда, когда $\mathbf{Af} = \mathbf{t}$. В этом случае система ограничений может рассматриваться как совокупность дополнительных наблю-

даемых полей, не пораженных шумом. При решении задач реставрации изображений можно ввести разумное ограничение, состоящее в том, что сумма значений элементов исправленного изображения должна равняться некоторому усредненному знечению, например среднему уровню наблюдаемого изображения. Однако такое ограничение оказывается не очень полезным для преодоления плохой обусловленности, так как влияет только на одну степень свсбоды в векторном пространстве размерности N^2.

Часто доступна априорная информация в виде ограничений-неравенств, обусловливающих значения элементов изображения. Физика процесса формирования изображений такова, что элементы изображения должны иметь неотрицательные значения. Часто известна верхняя граница этих значений, поскольку при преобразовании изображений в цифровую форму на каждый элемент отводится конечное число двоичных разрядов. Кроме того, вводятся вполне очевидные ограничения по уровню, связанные с необходимостью подогнать контрастный диапазон реставрированного изображения к динамическому диапазону дисплея. Один из возможных подходов — линейное масштабирование контраста исправленного изображения с учетом заданного динамического диапазона дисплея. Прибегать к этой процедуре обычно нежелательно, поскольку появление даже нескольких элементов с ненормально высокой яркостью приводит к снижению контраста всего изображения. Следует также учитывать, что контрастное масштабирование обычно влияет на среднюю яркость исправленного изображения. Другой распространенный метод воспроизведения изображений предусматривает ограничение уровня элементов в случае превышения порогового уровня дисплея. Хотя этот метод превосходит метод контрастного масштабирования по субъективному качеству изображения, он может дать смещение оценки.

Если реставрация изображений должна осуществляться с априорно вводимыми ограничениями уровней элементов, то лучше всего учесть эти ограничения непосредственно при реставрации; введение ограничений после завершения процедуры реставрации дает худший результат. Известно несколько методов реставрации изображений с использованием ограничений в виде неравенств.

Рассмотрим общий случай реставрации с ограничениями, когда вектор оценки $\hat{\mathbf{f}}$ должен удовлетворять неравенствам

$$\mathbf{l} \leqslant \hat{\mathbf{f}} \leqslant \mathbf{u}, \tag{14.8.5}$$

где $\mathbf{u}$ и $\mathbf{l}$ — векторы, учитывающие верхний и нижний пределы ограничения оценки элементов. В случае реставрации по критерию наименьших квадратов квадратическую ошибку следует ми-

нимизировать с учетом ограничения (14.8.5). Такой подход приводит к задаче квадратического программирования [19]. При использовании абсолютной ошибки задачу реставрации можно сформулировать как задачу линейного программирования [34,

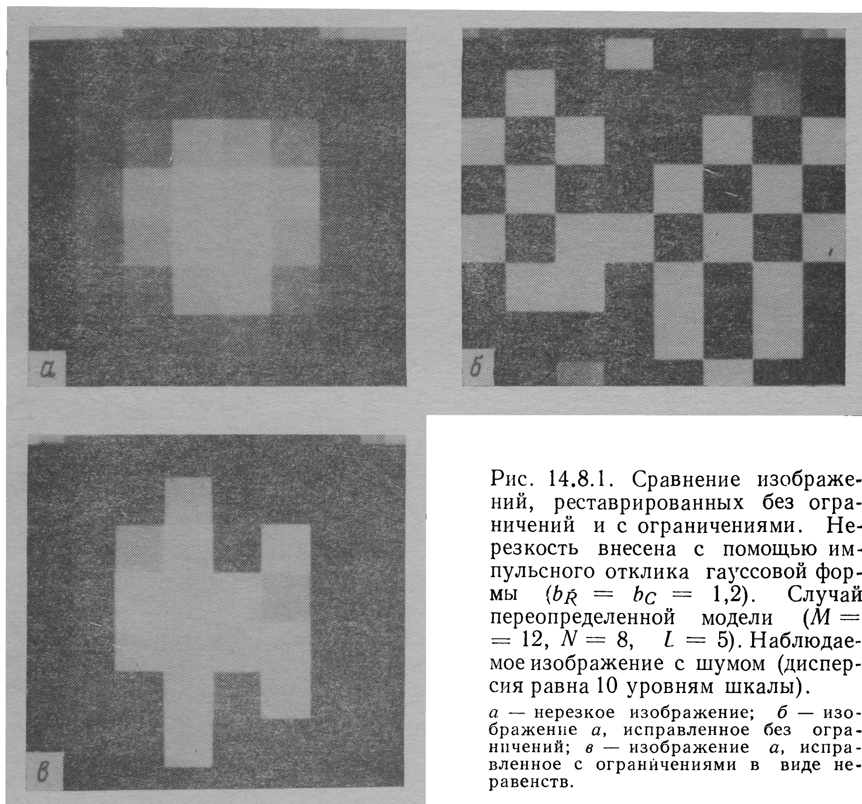


Рис. 14.8.1. Сравнение изображений, реставрированных без ограничений и с ограничениями. Нерезкость внесена с помощью импульсного отклика гауссовой формы ($b_R = b_C = 1{,}2$). Случай переопределенной модели ($M = 12$, $N = 8$, $L = 5$). Наблюдаемое изображение с шумом (дисперсия равна 10 уровням шкалы).

а — нерезкое изображение; *б* — изображение *а*, исправленное без ограничений; *в* — изображение *а*, исправленное с ограничениями в виде неравенств.

35]. Априорная информация, включающая ограничения в виде неравенств, может обеспечить существенное уменьшение неопределенности элементов исправленного изображения; однако, как и в случае использования ограничений-равенств, оценка может получить неизвестное смещение.

Рис. 14.8.1 иллюстрирует качество реставрации изображения с гауссовой нерезкостью, соответствующей переопределенной модели (см. гл. 13). Изображение на рис. 14.8.1, *б* реставрировано методом псевдообращения матриц, а изображение на рис. 14.8.1, *в* — с использованием ограничений-неравенств [19], требующих, чтобы яркость каждого элемента исправленного

изображения находилась в диапазоне 0—255. Введение ограничений обеспечивает существенное повышение качества реставрации. К сожалению, решение задачи квадратического программирования, которое использовано в рассматриваемом примере, требует выполнения большого объема вычислений. Распространить принцип «грубой силы» на рассматриваемый метод не представляется возможным.

Предложен ряд способов улучшения метода реставрации изображений с использованием ограничений. Один простой подход, основанный на идее гомоморфной фильтрации, заключается в логарифмировании каждого наблюдаемого изображения. Потенцирование соответствующих оценок автоматически приводит к строго положительному результату. Берг [5, 36, 37] и Фриден [5, 38, 39] разработали методы реставрации с ограничением в виде условия положительности, основанные на принципе максимума энтропии, который впервые был применен для оценивания плотности вероятности по ее моментам. Метод Берга, дающий решение в замкнутой форме, при наличии шума иногда приводит к неустойчивым оценкам. Итеративный метод Фридена с успехом был применен для обработки одномерных сигналов и небольших изображений. Джанссон и др. [40] на основе результатов более ранней работы Ван Циттера [41] разработали итеративный метод реставрации, в котором используются итерационные уравнения вида [5, 40]

$$\hat{\mathbf{g}}_k = \mathbf{B}\hat{\mathbf{f}}_k, \tag{14.8.6}$$

$$\hat{\mathbf{f}}_{k+1} = \hat{\mathbf{f}}_k + \mathbf{R}[\mathbf{g} - \hat{\mathbf{g}}_k], \tag{14.8.7}$$

где индекс указывает номер итерации, а $\hat{\mathbf{f}}_0 = \mathbf{g}$. Диагональная матрица $\mathbf{R}$ определяется как

$$R(n, n) = C\left[1 - \frac{2}{\mu(n) - l(n)}\left|\hat{f}_k(n) - \frac{l(n) + \mu(n)}{2}\right|\right], \tag{14.8.8}$$

где C — постоянная. Когда значение элемента исправленного изображения приближается к верхнему или нижнему пределу ограничения, соответствующий взвешивающий коэффициент $R(n, n)$ стремится к нулю, благодаря чему обеспечивается устойчивость этого члена оценки. Метод дал хорошие результаты при обработке одномерных сигналов. Хуанг и др. [42, 43] предложили модифицированный метод для реставрации изображений с ограничениями — так называемый метод проекций, предусматривающий итеративное решение системы уравнений $\mathbf{g} = \mathbf{B}\mathbf{f}$ численными методами. На каждом этапе промежуточные оценки подвергаются ограничению по уровню в соответствии с установленным динамическим диапазоном.

Реставрация изображений с ограничениями — развивающаяся область исследований. Здесь ведутся настойчивые поиски эффективных вычислительных методов, минимизирующих ошибки оценивания при наличии дополнительных ограничений.

14.9. МЕТОДЫ РЕСТАВРАЦИИ ИЗОБРАЖЕНИЙ НА ОСНОВЕ СТАТИСТИЧЕСКОГО ОЦЕНИВАНИЯ

Большинство из рассмотренных в этой главе методов реставрации изображений основано на использовании простой модели нерезкого изображения, пораженного аддитивным шумом. Более того, почти все эти методы разработаны для критериев наименьших квадратов или среднеквадратической ошибки. Теперь рассмотрим более общий класс методов реставрации изображений, основанных на теории статистического оценивания.

Обращаясь к формуле (13.7.2), запишем обобщенное векторное соотношение, связывающее вектор наблюдаемого поля $\mathbf{g}$ с векторами идеального изображения $\mathbf{f}$ и шума $\mathbf{n}$:

$$\mathbf{g} = \mathcal{O}_M\{\mathbf{f}, \mathbf{n}\}, \qquad (14.9.1)$$

где $\mathcal{O}_M(\cdot)$ — обобщенная передаточная функция, моделирующая поэлементные и пространственные искажения изображения. Для нахождения оценки $\hat{\mathbf{f}}$ по формуле (14.9.1) можно воспользоваться двумя общими подходами, основанными на групповой и рекуррентной обработке соответственно. Групповой процессор вычисляет оценку с использованием одновременно всех элементов наблюдаемого поля $\mathbf{g}$; напротив, рекуррентный процессор пользуется только теми элементами поля $\mathbf{g}$, которые расположены в некоторой окрестности оцениваемого элемента. При сильной корреляции элементов $\mathbf{g}$ групповой процессор вырабатывает более точную оценку, поскольку он использует больше входных данных. Ниже рассматриваются некоторые методы статистического оценивания [44—48], потенциально применимые к модели реставрации изображений (14.9.1).

Статистическая оценка самого общего вида минимизирует некоторый усредненный функционал ошибки, определенный как

$$\mathcal{E} = E\{\psi(\mathbf{f} - \hat{\mathbf{f}}) \mid \mathbf{g}\} = \int \psi(\mathbf{f} - \hat{\mathbf{f}})\, p(\mathbf{f} \mid \mathbf{g})\, d\mathbf{f}, \qquad (14.9.2)$$

где $\psi(\cdot)$ — функция ошибки или стоимостная функция, $p(\mathbf{f} \mid \mathbf{g})$ — апостериорная плотность вероятности вектора идеального изображения $\mathbf{f}$ при заданном наблюдаемом векторе $\mathbf{g}$. Если в качестве функционала ошибки используется среднеквадратическая ошибка

(14.6.2), получаемая оптимальная оценка называется бейесовской и имеет вид

$$\hat{\mathbf{f}} = E\{\mathbf{f}\,|\,\mathbf{g}\} = \int \mathbf{f} p(\mathbf{f}\,|\,\mathbf{g})\, d\mathbf{f}, \tag{14.9.3}$$

т. е. она является условным математическим ожиданием вектора идеального изображения при заданном наблюдаемом векторе. Бейесовская оценка будет оптимальной и в случае применения многих других функционалов ошибки (например, абсолютной ошибки) при условии, что функционал ошибки обладает симметрией, а апостериорная плотность вероятности симметрична и одномодальна [48]. К сожалению, в случае нелинейной модели (14.9.1) наблюдаемого поля вычисление бейесовской оценки (14.9.3) обычно связано с большими трудностями. Нараги [49] разработал алгоритм реставрации изображений, дающий приближение к оценке (14.9.3).

Другая распространенная статистическая оценка — так называемая оценка максимума апостериорной вероятности (МАВ), которая соответствует моде апостериорной плотности вероятности $p(\mathbf{f}\,|\,\mathbf{g})$. Для нахождения моды не обязательно располагать информацией об апостериорной плотности вероятности в явной форме. Теорема Бейеса позволяет выразить апостериорную плотность условной вероятности через плотности безусловных вероятностей $p(\mathbf{f})$ и $p(\mathbf{g})$ и плотность условной вероятности $p(\mathbf{g}\,|\,\mathbf{f})$:

$$p(\mathbf{f}\,|\,\mathbf{g}) = p(\mathbf{g}\,|\,\mathbf{f})\, p(\mathbf{f})/p(\mathbf{g}). \tag{14.9.4}$$

Логарифмируя обе части равенства (14.9.4) и дифференцируя результат по $\mathbf{f}$, получаем характеристическое уравнение вида

$$\partial \ln[p(\mathbf{g}\,|\,\mathbf{f})]/\partial\mathbf{f} + \partial \ln[p(\mathbf{f})]/\partial\mathbf{f} = 0, \tag{14.9.5}$$

которое требуется решить относительно оценки $\hat{\mathbf{f}}$ при наличии подходящих моделей для $p(\mathbf{g}\,|\,\mathbf{f})$ и $p(\mathbf{f})$. Большинство найденных решений основано на гауссовых моделях для $\mathbf{f}$ и $\mathbf{n}$. Хант [3] применил метод МАВ-оценки для решений задач реставрации изображений при использовании модели

$$\mathbf{g} = \mathcal{O}_M\{\mathbf{Bf}\} + \mathbf{n}, \tag{14.9.6}$$

где $\mathbf{B}$ — матрица нерезкости, $\mathcal{O}_M(\cdot)$ — известный оператор поэлементной нелинейности, воздействующий на каждый элемент вектора $\mathbf{Bf}$. Приняв гауссову модель, Хант представил (14.9.5) в рекуррентной форме, пригодной для вычисления решений итеративными методами.

Оценка наибольшего правдоподобия есть такое значение $\hat{\mathbf{f}}$, при котором апостериорная плотность вероятности $\mathbf{p}(\mathbf{g} = \tilde{\mathbf{g}}\,|\,\mathbf{f})$

для фактически наблюдаемого поля $\mathbf{g}$ максимальна. Математически такая оценка определяется как

$$p(\tilde{\mathbf{g}} \mid \hat{\mathbf{f}}) = \max_{\mathbf{f}} p(\tilde{\mathbf{g}} \mid \mathbf{f}). \qquad (14.9.7)$$

Если вектор идеального изображения $\mathbf{f}$ имеет равномерное распределение, оценка наибольшего правдоподобия совпадает с МАВ-оценкой.

Методы статистических оценок должны дать хорошие результаты при решении задач реставрации изображений с нелинейными моделями наблюдаемых полей изображения. В таких случаях методы линейных оценок представляются неэффективными. Однако статистические методы реставрации изображений, известные в настоящее время, находят ограниченное применение из-за трудностей статистического моделирования и сложности практической реализации.

Литература

1. Sondhi M. M., Image Restoration: The Removal of Spatially Invariant Degradations, *Proc. IEEE*, **60**, 7, 842—853 (July 1972). [Имеется перевод: Сондхи. Реставрация изображения: устранение пространственно-инвариантных искажений. — ТИИЭР, 1972, т. 60, № 7, с. 108—123.]
2. Andrews H. C., Digital Image Restoration: A Survey, *IEEE Computer*, **7**, 5, 36—45 (May 1974).
3. Hunt B. R., Digital Image Processing, *Proc. IEEE*, **63**, 4, 693—708 (April 1975). [Имеется перевод: Хант. Цифровая обработка изображений. — ТИИЭР, 1975, т. 63, с. 177—195.]
4. Andrews H. C., Hunt B. R., Digital Image Restoration, Prentice-Hall, Englewood Cliffs, N. J., 1977.
5. Frieden B. R., Image Enhancement and Restoration, in: Picture Processing and Digital Filtering, Huang T. S., Ed., Springer-Verlag, New York, 1975. [Имеется перевод: Фриден Б. Улучшение и реставрация изображения. — В кн.: Обработка изображений и цифровая фильтрация. [Под ред. Т. Хуанга. — М.: Мир, 1979.]
6. Merechal A., Croce P., Dietzel K., Amelioration du Contrast des Details des Images Photographiques par Filtrage des Frequencies Spatiales, *Opta Acta*, **5**, 256—262 (1958).
7. Tsujiuchi J., Correction of Optical Images by Compensation of Aberrations and by Spatial Frequency Filtering, in: Progress in Optics, Vol. 2, Wolf E., Ed., Wiley, New York, 1963, pp. 131—180.
8. Harris J. L., Sr., Image Evaluation and Restoration, *J. Opt. Soc. Am.*, **56**, 5, 569—574 (May 1966).
9. McGlamery B. L., Restoration of Turbulence-Degraded Images, *J. Opt. Soc. Am.*, **57**, 3, 293—297 (March 1967).
10. Mueller P. F., Reynolds G. O., Image Restoration by Removal of Random Media Degradations, *J. Opt. Soc. Am.*, **57**, 11, 1338—1344 (November 1967).
11. Helstrom C. W., Image Restoration by the Method of Least Squares, *J. Opt. Soc. Am.*, **57**, 3, 297—303 (March 1967).
12. Harris J. L., Sr., Potential and Limitations of Techniques for Processing Linear Motion-Degraded Imagery, in: Evaluation of Motion Degraded Images, U.S. Government Printing Office, Washington D. C., 1968, pp. 131—138.

13. Horner J. L., Optical Spatial Filtering with the Least-Mean-Square-Error Filter, *J. Opt. Soc. Am.*, **51**, 5, 553—558 (May 1969).
14. Horner J. L., Optical Restoration of Images Blurred by Atmospheric Turbulence Using Optimum Filter Theory, *Appl. Opt.*, **9**, 1, 167—171 (January 1970).
15. Lewis B. L., Sakrison D. J., Computer Enhancement of Scanning Electron Micrographs, *IEEE Trans. Circuits and Systems*, **CAS-22**, 3, 267—278 (March 1975).
16. Slepian D., Restoration of Photographs Blurred by Image Motion, *Bell Syst. Tech. J.*, **XLVI**, 10, 2353—2362 (December 1967).
17. Cole E. R., The Removal of Unknown Image Blurs by Homomorphic Filtering, Ph. D. Dissertation, Department of Electrical Engineering, University of Utah, Salt Lake City, June 1973.
18. Hunt B. R., The Application of Constrained Least Squares Estimation to Image Restoration by Digital Computer, *IEEE Trans. Computers*, **C-23**, 9, 805—812 (September 1973).
19. Mascarenhas N. D. A., Pratt W. K., Digital Image Restoration Under a Regression Model, *IEEE Trans. Circuits and Systems*, **CAS-22**, 3, 252—266 (March 1975).
20. Pratt W. K., Davarian F., Fast Computational Techniques for Pseudoinverse and Wiener Image Restoration, *IEEE Trans. Computers*, **C-26**, 6, 571—580 (June 1977).
21. Pratt W. K., Pseudoinverse Image Restoration Computational Algorithms, in: Optical Information Processing, Vol. II, Stroke G. W., Nesterikhin Y., Barrekette E. S., Eds., Plenum Press, New York, 1977.
22. Rust B. W., Burrus W. R., Mathematical Programming and the Numerical Solution of Linear Equations, American Elsevier, New York, 1972.
23. Albert A., Regression and the Moore-Penrose Pseudoinverse, Academic Press, New York, 1972.
24. Andrews H. C., Patterson C. L., Outer Product Expansions and Their Uses in Digital Image Processing, *Am. Math. Monthly*, **1**, 82, 1—13 (January 1975).
25. Andrews H. C., Patterson C. L., Outer Product Expansions and Their Uses in Digital Image Processing, *IEEE Trans. Computers*, **C-25**, 2, 140—148 (February 1976).
26. Huang T. S., Narendra P. M., Image Restoration by Singular Value Decomposition, *Appl. Opt.*, **14**, 9, 2213—2216 (September 1975).
27. Andrews H. C., Patterson C. L., Singular Value Decompositions and Digital Image Processing, *IEEE Trans. Acoustics, Speech, and Signal Processing*, **ASSP-24**, 1, 26—53 (February 1976).
28. Lewis T. O., Odell P. L., Estimation in Linear Models, Prentice Hall, Englewood Cliffs, N. J., 1971.
29. Pratt W. K., Generalized Wiener Filter Computation Techniques, *IEEE Trans. Computers*, **C-21**, 7, 636—641 (July 1972).
30. Papoulis A., Probability Random Variables and Stochastic Processes, McGraw-Hill, New York, 1965.
31. Twomey S., On the Numerical Solution of Fredholm Integral Equations of the First Kind by the Inversion of the Linear System Produced by Quadrature, *J. Assoc. Computing Machinery*, **10**, 97—101 (1963).
32. Phillips D. L., A Technique for the Numerical Solution of Certain Integral Equations of the First Kind, *J. Assoc. Computing Machinery*, **9**, 84—97 (1964).
33. Tikonow A. N., Regularization of Incorrectly Posed Problems, *Soviet Mathematics*, **4**, 6, 1624—1627 (1963).
34. Barrett E. B., Devich R. N., Linear Programming Compensation for Space-Variant Image Degradation, Processing SPIE/OSA Conference on Image Processing, Urbach J. C., Ed., Pacific Grove, California, February 1976, Vol. 74, pp. 152—158.

35. MacAdam D. P., Digital Image Restoration by Constrained Deconvolution. *J. Opt. Soc. Am.*, **60**, 12, 1617—1627 (December 1970).
36. Burg J. P., Maximum Entropy Spectral Analysis, 37th Annual Society of Exploration Geophysicists Meeting, Oklahoma City, 1967.
37. Edward J. A., Fitelson M. M., Notes on Maximum Entropy Processing, *IEEE Trans. Inf. Theory*, **IT-19**, 2, 232—234 (March 1973).
38. Frieden B. R., Restoring with Maximum Likelihood and Maximum Entropy, *J. Opt. Soc. Am.*, **62**, 4, 511—518 (April 1972).
39. Frieden B. R., Maximum Entropy Restorations of Garrymede, in: Proceedings SPIE/OSA Conference on Image Processing, Urbach J. C., Ed., Pacific Grove, California, February 1976, Vol. 74, pp. 160—165.
40. Jansson P. A., Hunt R. H., Peyler E. K., Resolution Enhancement of Spectra. *J. Opt. Soc. Am.*, **60**, 5, 596—599 (May 1970).
41. Van Cittert P. H., Zun einflusz der Spaltbreite auf die Intensitatsverteilung in spektrallinien-Part II, *Z. Physik*, **69**, 298—308 (1931).
42. Huang T. S., Baker D. S., Berger S. P., Iterative Image Restoration, *Appl. Opt.*, **14**, 5, 1165—1168 (May 1975).
43. Huang T. S., Restoring Images with Shift-Varying Degradations, Proceedings SPIE/OSA Conference on Image Processing, Urbach J. C., Ed., Pacific Grove, California, February 1976, Vol. 74, pp. 149—151.
44. Helstrom C. W., Statistical Theory of Signal Detection, Pergamon Press, New York, 1960. [Имеется перевод: Хелстром К., Статистическая теория обнаружения сигналов. — М.: ИЛ, 1963.]
45. Middleton D., An Introduction to Statistical Communication Theory, McGraw-Hill, New York, 1960. [Имеется перевод: Миддлтон Д., Введение в статистическую теорию связи. — М.: Советское радио, т. 1, 1961, т. 2, 1962.]
46. Deutsch R., Estimation Theory, Prentice-Hall, Englewood Cliffs, N. J., 1965.
47. Van Trees H. L., Detection, Estimation, and Modulation Theory, Vol. 1, Wiley, New York, 1968. [Имеется перевод: Ван Трис Г., Теория обнаружения, оценок и модуляции. т. 1. Теория обнаружения, оценок и линейной модуляции. — М.: Советское радио, 1972.]
48. Nahi N. E., Estimation Theory and Applications, Wiley, New York, 1969.
49. Naraghi M., An Algorithmic Image Estimation Method Applicable to Nonlinear Observations, University of Southern California, Image Processing Institute, USCIPI Report 580, 1975.

Глава 15

СПЕЦИАЛЬНЫЕ МЕТОДЫ ПРОСТРАНСТВЕННОЙ РЕСТАВРАЦИИ ИЗОБРАЖЕНИЙ

Гл. 14 посвящена рассмотрению основных методов пространственной реставрации изображений. В данной главе описывается применение этих методов для решения ряда конкретных задач.

15.1. РЕСТАВРАЦИЯ ИЗОБРАЖЕНИЙ С ЦЕЛЬЮ ПОДАВЛЕНИЯ ОШИБОК НАЛОЖЕНИЯ СПЕКТРОВ

Непрерывное изображение можно дискретизировать с пространственной частотой, составляющей меньше половины его наибольшей пространственной частоты. В этом случае будет формироваться массив отсчетов, содержащий ложные пространственные составляющие, которые возникают вследствие наложения на исходный спектр побочных спектров дискретизованного изображения. Процедура реставрации изображений может уменьшить вредное влияние этих ошибок [1, 2].

На рис. 15.1.1 показана модель системы с дискретизацией и восстановлением изображения. Пусть исходное изображение, описываемое функцией $F_I(x, y)$, дискретизируется с частотой ω_{xs} по координате x и ω_{ys} по координате y. Используются дискретизирующие импульсы в виде бесконечной последовательности дельта-функций, однако частоты дискретизации меньше частоты дискретизации Найквиста. Если бы дискретизация выполнялась в соответствии с критерием Найквиста, на выходе идеального фильтра нижних частот с областью пропускания $\pm\omega_{xs}/2$ и $\omega_{ys}/2$ формировалось изображение, идентичное исходному изображению. В случае дискретизации с пониженной частотой на выходе интерполирующего фильтра нижних частот возникает интерполированное изображение, описываемое функцией

$$F_R(x, y) = F_I(x, y) + A(x, y), \tag{15.1.1}$$

где $A(x, y)$ — «изображение» из ложных составляющих, вызванных наложением спектров. Согласно формуле (4.2.16), это «изображение» описывается функцией

$$A(x, y) = (1/4\pi^2) \int_{-\omega_{xs}/2}^{\omega_{xs}/2} \int_{-\omega_{ys}/2}^{\omega_{ys}/2} \mathcal{F}_Q(\omega_x, \omega_y) \exp\{i[\omega_x x + \omega_y y]\}\, d\omega_x\, d\omega_y, \tag{15.1.2}$$

где

$$\mathcal{F}_Q(\omega_x, \omega_y) = \sum_{\substack{j_1=-\infty \\ j_1,\, j_2 \neq 0}}^{\infty} \sum_{j_2=-\infty}^{\infty} \mathcal{F}_I(\omega_x - j_1\omega_{xs}, \omega_y - j_2\omega_{ys}) \quad (15.1.3)$$

представляет составляющие побочных спектров дискретизованного изображения.

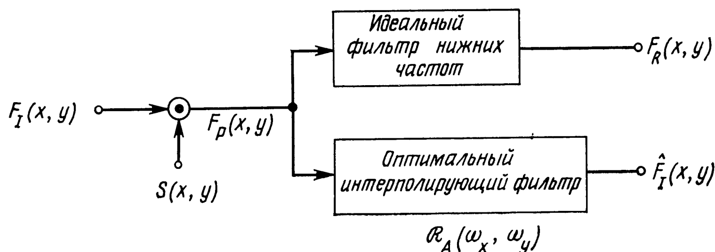


Рис. 15.1.1. Модель реставрации изображений с целью подавления ложных частот.

Выражение (15.1.1) соответствует классической модели сигнала с помехой; следовательно, борьбу с ложными пространственными частотами можно вести, пользуясь классическими методами реставрации изображений, описанными в гл. 14. Если исходное изображение, описываемое функцией $F_I(x, y)$, рассматривать как реализацию двумерного случайного процесса, то $A(x, y)$ также будет случайным полем. В этом случае частотную характеристику оптимального интерполирующего фильтра, минимизирующего среднеквадратическое отклонение оценки $\hat{F}_I$ от исходного изображения F_I, можно найти методами винеровского оценивания. Если речь идет о непрерывных изображениях, то оптимальная интерполирующая частотная характеристика, определенная в области пропускания фильтра нижних частот, имеет вид

$$\mathcal{R}_A(\omega_x, \omega_y) = \mathcal{W}_{F_I F_R}(\omega_x, \omega_y) / \mathcal{W}_{F_R}(\omega_x, \omega_y). \quad (15.1.4)$$

В этом выражении числитель есть взаимный энергетический спектр F_I и интерполированного изображения F_R, знаменатель — энергетический спектр F_R. При относительной независимости высокочастотных и низкочастотных пространственных составляющих элементы $F_I(x, y)$ и $A(x, y)$ могут считаться независимыми. В рамках этого допущения частотная характеристика интерпо-

лирующего фильтра в области пропускания описывается выражением

$$\mathscr{R}_A(\omega_x, \omega_y) = \mathscr{W}_{F_I}(\omega_x, \omega_y)/[\mathscr{W}_{F_I}(\omega_x, \omega_y) + \mathscr{W}_A(\omega_x, \omega_y)], \quad (15.1.5)$$

где $\mathscr{W}_A(\omega_x, \omega_y)$ — энергетический спектр ложных составляющих $A(x, y)$.

15.2. КОРРЕКЦИЯ ОШИБОК ИНТЕРПОЛЯЦИИ

В большинстве воспроизводящих изображения дисплеев затруднительно формировать и поддерживать с большой точностью параметры апертуры (пятна). Следовательно, на практике редко удается обеспечить оптимальную интерполяцию отсчетов дискретизованного изображения. Однако во многих случаях ошибки интерполяции, вносимые дисплеем, удается скомпенсировать посредством цифровой обработки. Реставрации поддаются также многие плохо интерполированные изображения, записанные фотографическими или электронными методами.

В разд. 13.5 приведены некоторые модели интерполяции, производимой при воспроизведении изображений. Интерполированное непрерывное изображение можно описать функцией

$$F_R(x, y) = \sum_{k_1=-K}^{K} \sum_{k_2=-K}^{K} F_G(k_1, k_2) R(x - k_1 \Delta S, y - k_2 \Delta S), \quad (15.2.1)$$

где $F_G(k_1, k_2)$ — массив отсчетов изображения, ΔS — шаг дискретизации, $R(x, y)$ — интерполирующая функция дисплея, которая предполагается пространственно-инвариантной. Желаемое интерполированное изображение описывается функцией

$$F_D(x, y) = \sum_{k_1=-K}^{K} \sum_{k_2=-K}^{K} F_G(k_1, k_2) R_D(x - k_1 \Delta S, y - k_2 \Delta S), \quad (15.2.2)$$

где $R_D(x, y)$ — желаемая интерполирующая функция. Если R_D образована двумя двумерными sinc-функциями, F_D будет идеальным представлением исходного изображения F_I (если не считать ошибки усечения вследствие конечности разложения); при этом предполагается, что массив отсчетов F_G был получен с использованием частоты дискретизации Найквиста.

Непрерывные поля F_R и F_D можно сравнить, изучая их соответствующие отсчеты, расположенные в узлах решетки с шагом ΔI. Согласно выражениям (13.5.4) и (13.5.5), обозначим через

$$\mathbf{f}_R = \mathbf{R}\mathbf{f}_G, \quad (15.2.3)$$

$$\mathbf{f}_D = \mathbf{R}_D\mathbf{f}_G \quad (15.2.4)$$

векторные представления получаемых массивов отсчетов, где $\mathbf{R}$ и $\mathbf{R}_D$ — матрицы, элементами которых являются отсчеты $R(x, y)$ и $R_D(x, y)$ соответственно. Например, при использовании одномерной интерполирующей функции с пятью отсчетами и $\Delta S = \Delta I$ интерполирующая матрица имеет вид

$$\mathbf{R} = \begin{bmatrix} r(3) & r(2) & r(1) & 0 & \dots & 0 & 0 \\ r(4) & r(3) & r(2) & r(1) & & . & . \\ r(5) & r(4) & r(3) & r(2) & & . & . \\ 0 & r(5) & r(4) & r(3) & & . & . \\ . & & r(5) & r(4) & & . & . \\ . & & & r(5) & & . & . \\ . & & & & & . & . \\ . & & & & & . & . \\ . & & & & & . & . \\ . & & & & & r(1) & 0 \\ . & & & & & r(2) & r(1) \\ . & & & & & r(3) & r(2) \\ 0 & 0 & 0 & 0 & \dots & r(4) & r(3) \end{bmatrix}. \quad (15.2.5)$$

Каждый столбец $\mathbf{R}$ содержит значения отсчетов $r(x)$, расположенных с шагом ΔI.

Векторные соотношения (15.2.3) и (15.2.4) указывают на принципиальную возможность исправления ошибок, обусловленных несовершенством системы интерполирующих функций $\mathbf{R}$, посредством линейного преобразования массива отсчетов $\mathbf{f}_G$ с помощью оператора $\mathbf{W}$ перед интерполяцией. Ошибку по критерию наименьших квадратов, выражающую различие между $\mathbf{f}_R$ и $\mathbf{f}_D$, минимизирует оператор обобщенного обращения $\mathbf{W} = \mathbf{R}^- \mathbf{R}_D$, где $\mathbf{R}^-$ — обобщенная обратная матрица для $\mathbf{R}$. Преобразование без ошибок возможно лишь в том случае, когда число строк матрицы $\mathbf{R}$ меньше или равно числу ее столбцов. Если интерполированное поле $F_R(x, y)$ измеряется только в точках, разделенных интервалом Найквиста ΔS, то матрица $\mathbf{R}$ является квадратной, а оператор $\mathbf{W} = \mathbf{R}^- \mathbf{R}_D$ обеспечивает идеальную интерполяцию в точках расположения отсчетов изображения при условии существования $\mathbf{R}^{-1}$. Результаты измерений на интервалах между отсчетами предсказать невозможно.

Чрезвычайно большой размер матрицы $\mathbf{R}$ делает практически невозможным ее прямое обращение или псевдообращение. В случае квадратных матриц это затруднение можно обойти, сделав

$\mathbf{R}$ блочно-циклической матрицей [3]. В рассмотренном одномерном примере матрицу (15.2.5) легко обратить в циклическую матрицу

$$\tilde{\mathbf{R}} = \mathbf{R} + \begin{bmatrix} 0 & 0 & \dots & 0 & r(3) & r(4) \\ . & . & & 0 & 0 & r(5) \\ . & . & & 0 & 0 & 0 \\ . & . & & . & . & . \\ . & . & & . & . & . \\ r(1) & 0 & & . & . & . \\ r(2) & r(1) & \dots & 0 & 0 & 0 \end{bmatrix}. \qquad (15.2.6)$$

Таким же образом можно сформировать циклическую матрицу $\tilde{\mathbf{R}}_D$ из матрицы $\mathbf{R}_D$. Поскольку $\tilde{\mathbf{R}}$ и $\tilde{\mathbf{R}}_D$ — циклические матрицы, оператор $\tilde{\mathbf{W}} = \tilde{\mathbf{R}}^{-1}\tilde{\mathbf{R}}_D$ также является циклическим. Это позволяет воспользоваться методами, описанными в гл. 11, и осуществить реставрацию в фурье-области.

15.3. КОРРЕКЦИЯ ГЕОМЕТРИЧЕСКИХ ИСКАЖЕНИЙ

Многие изображающие системы вносят геометрические искажения. Рис. 15.3.1 иллюстрирует подушкообразную и бочкообразную дисторсию, часто возникающую в видиконных теле-

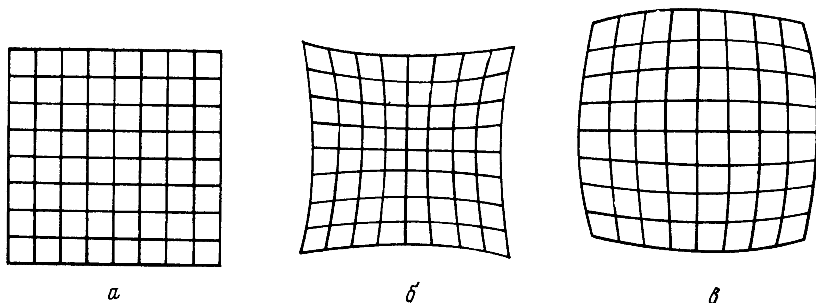


Рис. 15.3.1. Примеры геометрических искажений изображения.
а — искажения отсутствуют; *б* — подушкообразная дисторсия; *в* — бочкообразная дисторсия.

камерах и дисплеях с кинескопом. Один из возможных способов коррекции подобных искажений сводится к такому нелинейному предыскажению растра при развертке изображения, которое обеспечивает компенсацию ожидаемых искажений. Возможна также апостериорная коррекция, основанная на полиномиальной аппроксимации каждой горизонтальной и вертикальной линии иска-

женного растра; данные аппроксимации используются затем для вычисления обратных корректирующих функций для каждой ячейки сетки [4]. Метод пространственного искривления позволяет также корректировать перспективные искажения. Изображение протяженного объекта, наблюдаемого сбоку, можно искривить таким образом, что будет получено изображение, соответст-

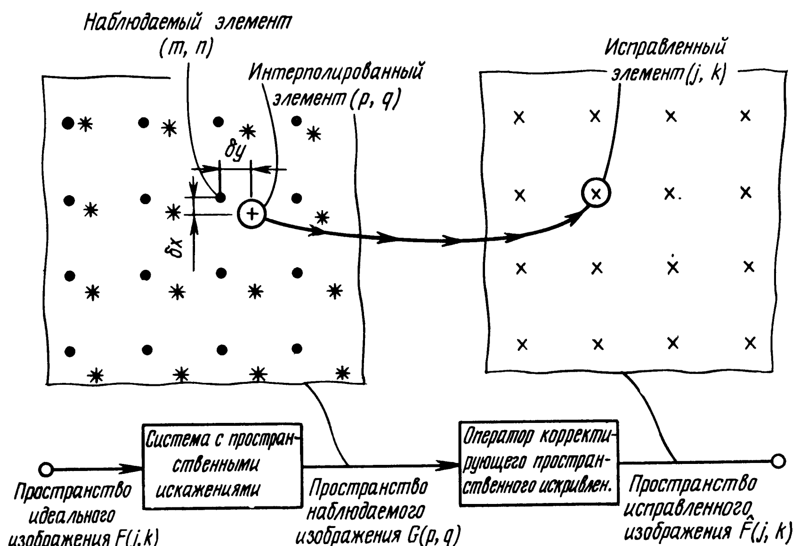


Рис. 15.3.2. Коррекция геометрических искажений методом пространственного искривления.

вующее наблюдению этого объекта под прямым углом. Другое важное применение этого метода — спрямление парных изображений одной и той же сцены, полученных при разных углах наблюдения.

Рис. 15.3.2 иллюстрирует идею пространственного искривления как метода коррекции пространственных искажений, вносимых изображающей системой. В идеальное изображение вносится пространственное геометрическое искажение, которое можно описать отображением точки $(j,\ k)$ идеального изображения в точку $(p,\ q)$ наблюдаемого изображения в соответствии с выражениями

$$p = \mathcal{O}_p\{j,\ k\}, \tag{15.3.1а}$$

$$q = \mathcal{O}_q\{j,\ k\}, \tag{15.3.1б}$$

где $\mathcal{O}_p\ \{\cdot\}$ и $\mathcal{O}_q\ \{\cdot\}$ — монотонные отображающие функции общего вида. На рис. 15.3.2 точками показаны элементы наблюдаемого изображения, соответствующие целочисленным значениям

координат точки (p, q). Скорректированные элементы изображения, показанные крестами, получены в результате корректирующего пространственного искривления (отображения) массива наблюдаемых элементов изображения, помеченных звездочками; в общем случае значения элементов со звездочкой не совпадают со значениями наблюдаемых элементов. Процедуру корректирующего искривления можно разбить на два этапа. Сначала определяют координаты точки (p, q), которая отображается в точку размещения скорректированного элемента изображения. Затем находят оценку яркости изображения в точке (p, q), интерполируя значения яркости соседних наблюдаемых элементов, и приписывают полученную оценку скорректированному элементу изображения.

Если функции пространственных искажений $\mathcal{O}_p\{\cdot\}$ и $\mathcal{O}_q\{\cdot\}$ известны, то координаты (p, q) можно вычислить непосредственно по формулам (15.3.1), подставляя в них координаты (j, k), где j и k — целые числа. Как правило, такая информация отсутствует, поэтому необходимо пользоваться некоторой математической моделью реальных искажений. Часто применяют полиномиальные модели. В качестве примера рассмотрим оценку пространственного искажения в виде полинома второго порядка:

$$\hat{p} = a_0 + a_1 j + a_2 k + a_3 j^2 + a_4 jk + a_5 k^2, \tag{15.3.2а}$$

$$\hat{q} = b_0 + b_1 j + b_2 k + b_3 j^2 + b_4 jk + b_5 k^2, \tag{15.3.2б}$$

где a_i и b_i — постоянные коэффициенты. Эти коэффициенты обычно выбирают таким образом, чтобы минимизировать среднеквадратическую ошибку аппроксимации фактически наблюдаемых координат (p_m, q_m) их полиномиальной оценкой $(\hat{p}_m, \hat{q}_m)$ для набора известных точек (j_m, k_m), называемых узлами $(1 \leqslant m \leqslant M)$. Координаты в плоскости наблюдаемого изображения удобно выразить в виде векторов

$$\mathbf{p}^T = [p_1, p_2, \ldots, p_M], \tag{15.3.3а}$$

$$\mathbf{q}^T = [q_1, q_2, \ldots, q_M]. \tag{15.3.3б}$$

Аналогично коэффициенты полиномов можно представить в векторной форме:

$$\mathbf{a}^T = [a_0, a_1, \ldots, a_5], \tag{15.3.4а}$$

$$\mathbf{b}^T = [b_0, b_1, \ldots, b_5]. \tag{15.3.4б}$$

Теперь среднеквадратическую ошибку оценивания можно записать в компактной форме:

$$\mathscr{E} = (\mathbf{p} - \mathbf{A}\mathbf{a})^T (\mathbf{p} - \mathbf{A}\mathbf{a}) + (\mathbf{q} - \mathbf{A}\mathbf{b})^T (\mathbf{q} - \mathbf{A}\mathbf{b}), \quad (15.3.5)$$

где

$$\mathbf{A} = \begin{bmatrix} 1 & j_1 & k_1 & j_1^2 & j_1k_1 & k_1^2 \\ 1 & j_2 & k_2 & j_2^2 & j_2k_2 & k_2^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & j_M & k_M & j_M^2 & j_Mk_M & k_M^2 \end{bmatrix}. \quad (15.3.6)$$

В гл. 8 было показано, что ошибка минимальна при

$$\mathbf{a} = \mathbf{A}^{-}\mathbf{p}, \quad (15.3.7а)$$

$$\mathbf{b} = \mathbf{A}^{-}\mathbf{q}, \quad (15.3.7б)$$

где $\mathbf{A}^{-}$ — обобщенная обратная матрица для $\mathbf{A}$. Если число узлов превышает число полиномиальных коэффициентов, то

$$\mathbf{A}^{-} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T. \quad (15.3.8)$$

Пользуясь этой методикой, можно без труда вычислить коэффициенты a_i, b_i, а затем по формулам (15.3.2) определить в плоскости наблюдаемого изображения координаты всех точек (звездочки на рис. 15.3.2), отображаемых в точки с целочисленными координатами, которые соответствуют скорректированным элементам изображения (кресты на рис. 15.3.2).

Теперь, зная координаты (p, q) в плоскости наблюдаемого изображения и уровни яркости наблюдаемых элементов, нужно оценить уровень яркости элемента с координатами (j, k) в плоскости скорректированного изображения. Эту задачу можно решить методом интерполяции. Согласно формуле (4.1.9), интерполированное непрерывное изображение в плоскости наблюдаемого изображения можно описать функцией

$$G(x, y) = \sum_m \sum_n G(m\Delta, n\Delta) R(x - m\Delta, y - n\Delta), \quad (15.3.9)$$

где $R(x, y)$ — интерполирующая функция, Δ — шаг дискретизации. Оценка скорректированного изображения запишется как

$$\hat{F}(j, k) = G(p\Delta, q\Delta), \quad (15.3.10)$$

где (p, q) — нецелочисленный индекс интерполированного элемента. Теоретически оптимальную интерполяцию обеспечивают sinc-функции или функции Бесселя. На практике, однако, обычно пользуются интерполирующими функциями, реализация которых сопряжена с меньшими трудностями. Это прямоугольные, треугольные и В-сплайн функции, рассмотренные в разд. 4.3. В некоторых системах с пространственным искривлением используют даже интерполяцию по окрестности элемента изображения.

15.4. РЕСТАВРАЦИЯ ИЗОБРАЖЕНИЙ МЕТОДОМ ГОМОМОРФНОЙ ФИЛЬТРАЦИИ

Если наблюдаемое изображение искажено мультипликативной помехой, его целесообразно реставрировать методом гомоморфной фильтрации [5], суть которой иллюстрирует рис. 15.4.1. Пусть

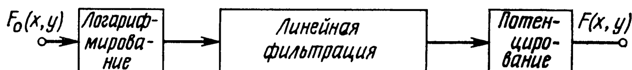


Рис. 15.4.1. Реставрация изображений методом гомоморфной фильтрации.

наблюдаемое непрерывное изображение представляется функцией, полученной в результате умножения функции, описывающей идеальное изображение, на функцию $E(x, y)$:

$$F_O(x, y) = F_I(x, y) E(x, y). \qquad (15.4.1)$$

Функция $E(x, y)$ может описывать мультипликативный шум или, например, неравномерность освещенности по полю. В любом случае цель реставрации состоит в нахождении оценки $F_I(x, y)$ по данным обработки наблюдаемого изображения. Логарифмирование соотношения (15.4.1) приводит к аддитивным помехам:

$$\log [F_O(x, y)] = \log [F_I(x, y)] + \log [E(x, y)]. \qquad (15.4.2)$$

Оценку логарифма $F_I(x, y)$ можно найти, пользуясь обычными методами линейной фильтрации. Последующее антилогарифмирование дает оценку идеального изображения.

15.5. СЛЕПАЯ РЕСТАВРАЦИЯ ИЗОБРАЖЕНИЙ

Почти все методы реставрации изображений основаны на использовании априорной информации об искажении, внесенном в изображение; предполагается, что известны поэлементная характеристика передачи уровней и пространственный импульсный отклик. Во многих применениях такая информация недоступна. Может оказаться, что искажение либо трудно измерить, либо оно претерпевает непредсказуемые изменения во времени. Во всех таких случаях информацию об искажениях приходится извлекать из наблюдаемого изображения, пользуясь явными или неявными методами. Реставрация изображений в таких условиях называется слепой [6—10]. В данной главе рассмотрение методов слепой реставрации изображений ограничено вопросами обработки нерезкого изображения с аддитивным шумом.

Известны два основных подхода к решению задачи слепой реставрации изображений: прямое измерение и косвенное оценивание. Применяя методы первой группы, прежде всего оценивают

импульсный отклик нерезкости и уровень шума на основе измеренных параметров реставрируемого изображения; полученная информация используется затем при реставрации изображения. Методы косвенного оценивания основаны на процедуре временного или пространственного усреднения, причем получаемые данные используют либо для осуществления реставрации изображения, либо для определения ключевых элементов реставрационного алгоритма.

15.5.1. МЕТОДЫ НА ОСНОВЕ ПРЯМЫХ ИЗМЕРЕНИЙ

Слепая реставрация нерезкого зашумленного изображения обычно сопряжена с измерением импульсного отклика нерезкости и энергетического спектра шума или ковариационной функции наблюдаемого изображения. Импульсный отклик нерезкости обычно измеряют, изолируя участок изображения, соответствующий предположительно известному объекту. По определению импульсный отклик есть изображение точечного источника. Следовательно, любой точечный объект наблюдаемого изображения предоставляет прямую информацию об импульсном отклике. Изображение предположительно резкого перепада яркости также может служить основой для определения импульсного отклика. Усреднение нескольких строк изображения, перпендикулярных направлению перепада, обеспечивает существенное снижение влияния шума. Ковариационную функцию шума наблюдаемого изображения удается оценить посредством измерения ковариационной функции изображения на участке с относительно постоянной фоновой яркостью. В работах [6—8] содержится подробное описание методов реставрации изображений на основе прямых измерений.

15.5.2. МЕТОДЫ НА ОСНОВЕ КОСВЕННОГО ОЦЕНИВАНИЯ

Для телевизионных систем, действующих в реальном масштабе времени, характерна временна́я избыточность передаваемых кадров изображения, которая позволяет реализовать слепую реставрацию косвенными методами. Пусть в телевизионной системе i-й кадр наблюдаемого изображения описывается функцией

$$G_i(x, y) = F_I(x, y) + N_i(x, y), \qquad (15.5.1)$$

где $F_I(x, y)$ — функция идеального изображения, $N_i(x, y)$ — поле аддитивного шума, не коррелированного с $F_I(x, y)$. Если на M кадрах идеальное изображение сохраняется неизменным, то временно́е усреднение наблюдаемых полей дает следующий результат:

$$F_I(x, y) = (1/M) \sum_{i=1}^{M} G_i(x, y) - (1/M) \sum_{i=1}^{M} N_i(x, y). \qquad (15.5.2)$$

При большом M величина отрицательного члена (15.5.2) стремится к среднему по ансамблю значению шума $E\{N(x, y)\}$. В распространенном случае гауссового белого шума с нулевым средним для всех (x, y) имеем нулевое среднее по ансамблю; это обстоятельство дает основание записать оценку в виде

$$\hat{F}_I(x, y) = (1/M) \sum_{i=1}^{M} G_i(x, y). \qquad (15.5.3)$$

Временнóе усреднение также позволяет компенсировать нерезкость изображения. Рассмотрим изображающую систему, в которой последовательные кадры содержат изображения почти неподвижных объектов, причем в каждом кадре нерезкость вносится импульсным откликом $H_i(x, y)$, независимым от линейного сдвига. Примером такой изображающей системы может служить фотографическая система, в которой удаленный объект наблюдают через турбулентную атмосферу; при этом предполагается, что в межкадровых интервалах смещение объекта несущественно. Съемка с короткой выдержкой позволяет «заморозить» пространственную картину атмосферной турбулентности в моменты времени, соответствующие последовательным кадрам изображения. Для i-го кадра нерезкого изображения такого объекта можно записать

$$G_i(x, y) = F_I(x, y) * H_i(x, y), \qquad (15.5.4)$$

где $i = 1, 2, \ldots, M$. Фурье-спектр искаженного изображения запишется так

$$\mathcal{G}_i(\omega_x, \omega_y) = \mathcal{F}_I(\omega_x, \omega_y)\mathcal{H}_i(\omega_x, \omega_y). \qquad (15.5.5)$$

Логарифмирование выражения (15.5.5) дает

$$\ln[\mathcal{G}_i(\omega_x, \omega_y)] = \ln[\mathcal{F}_I(\omega_x, \omega_y)] + \ln[\mathcal{H}_i(\omega_x, \omega_y)], \qquad (15.5.6)$$

т. е. делает аддитивно разделимыми спектр идеального изображения и передаточную функцию искажающей системы. Это позволяет обратиться к общеизвестным методам статистического оценивания сигнала на фоне аддитивного шума. При отсутствии корреляции между импульсными откликами отдельных кадров изображения целесообразно вычислить сумму

$$\sum_{i=1}^{M} \ln[\mathcal{G}_i(\omega_x, \omega_y)] = M \ln[\mathcal{F}_I(\omega_x, \omega_y)] + \sum_{i=1}^{M} \ln[\mathcal{H}_i(\omega_x, \omega_y)], \qquad (15.5.7)$$

поскольку при большом M она стремится к постоянной величине

$$\mathcal{K}_H(\omega_x, \omega_y) = \lim_{M\to\infty} \sum_{i=1}^{M} \ln[\mathcal{H}_i(\omega_x, \omega_y)]. \qquad (15.5.8)$$

Таким образом, $\mathcal{K}_H(\omega_x, \omega_y)$ можно рассматривать как результат усреднения логарифмов частотных характеристик, описывающих турбулентность атмосферы. Выражение для оценки спектра изображения можно записать в виде

$$\hat{\mathcal{F}}_I(\omega_x, \omega_y) = \exp\{-\mathcal{K}_H(\omega_x, \omega_y)/M\} \prod_{i=1}^{M} [\mathcal{G}_i(\omega_x, \omega_y)]^{1/M}. \quad (15.5.9)$$

Обратное преобразование Фурье-спектра (15.5.9) дает оценку в пространственной области. В любой практической изображающей системе в формулу (15.5.4) приходится вводить член для учета аддитивного шума $N_i(x, y)$. К сожалению, эта шумовая составляющая делает неправомерной операцию разделения (15.5.6) и, следовательно, заставляет отказаться от всех последующих выкладок. Один из методов преодоления этой трудности заключается в предварительном сглаживании (фильтрации) шума каждого наблюдаемого поля и использовании полученных результатов в формуле (15.5.9) в качестве оценок «незашумленных» изображений [9]. Можно также применить метод слепой реставрации сигналов, разработанный Стокхэмом и др. [9, 11].

Этот метод можно использовать, в частности, для обработки одиночных кадров зашумленных нерезких изображений, описываемых как

$$F_O(x, y) = F_I(x, y) * H_D(x, y) + N(x, y), \quad (15.5.10)$$

где $H_D(x, y)$ — пространственно-инвариантный импульсный отклик нерезкости. Реставрация выполняется с помощью сохраняющего энергетический спектр изображения фильтра с частотной характеристикой (14.1.16):

$$\mathcal{H}_R(\omega_x, \omega_y) = \left[\frac{\mathcal{W}_{F_I}(\omega_x, \omega_y)}{|\mathcal{H}_D(\omega_x, \omega_y)|^2 \mathcal{W}_{F_I}(\omega_x, \omega_y) + \mathcal{W}_N(\omega_x, \omega_y)}\right]^{1/2}, \quad (15.5.11)$$

где $\mathcal{W}_{F_I}(\omega_x, \omega_y)$ и $\mathcal{W}_N(\omega_x, \omega_y)$ — энергетические спектры идеального изображения и шума соответственно, $\mathcal{H}_D(\omega_x, \omega_y)$ — частотная характеристика нерезкости. При использовании метода слепой реставрации изображений Стокхэма оценку знаменателя в (15.5.11) находят, разбивая наблюдаемое изображение на фрагменты из 64×64 элементов, которые можно представить как

$$\tilde{F}_{Oi}(x, y) = \tilde{F}_{Ii}(x, y) * H_D(x, y) + \tilde{N}_i(x, y). \quad (15.5.12)$$

Каждый фрагмент обрабатывается соответствующим окном и подвергается преобразованию Фурье:

$$\tilde{\mathscr{F}}_{Oi}(\omega_x, \omega_y) = \tilde{\mathscr{F}}_{Ii}(\omega_x, \omega_y)\mathscr{H}_D(\omega_x, \omega_y) + \tilde{\mathscr{N}}_i(\omega_x, \omega_y). \tag{15.5.13}$$

Затем для каждого фрагмента вычисляют логарифм модуля спектра, а полученные данные усредняют. В результате получают

$$\mathscr{L} = \mathscr{R}, \tag{15.5.14}$$

где

$$\mathscr{L} \equiv (1/N_B) \sum_{i=1}^{N_B} \log |\tilde{\mathscr{F}}_{Oi}(\omega_x, \omega_y)|, \tag{15.5.15а}$$

$$\mathscr{R} = (1/N_B) \sum_{i=1}^{N_B} \log |\tilde{\mathscr{F}}_{Ii}(\omega_x, \omega_y)\mathscr{H}_D(\omega_x, \omega_y) + \mathscr{N}(\omega_x, \omega_y)|, \tag{15.5.15б}$$

где N_B — число фрагментов. Если N_B достаточно велико, пространственные средние в обеих частях равенства (15.5.14) мало отличаются от соответствующих средних по ансамблю. Было показано [9], что среднее и дисперсия для (15.5.15б) имеют вид

$$\overline{\mathscr{R}} = {}^1/_2 \log [\mathscr{W}_{F_I}(\omega_x, \omega_y)\,|\mathscr{H}_D(\omega_x, \omega_y)|^2 + \mathscr{N}(\omega_x, \omega_y)] + (C/2), \tag{15.5.16а}$$

$$\sigma^2_{\mathscr{R}} = \pi^2/24N_B, \tag{15.5.16б}$$

где $C = 0{,}57221 \ldots$ — постоянная Эйлера. Таким образом, равенство $\mathscr{L} = \overline{\mathscr{R}}$ позволяет найти оценку знаменателя в (15.5.11) для частотной характеристики реставрирующего фильтра. Оценку числителя формулы (15.5.11) находят подобным образом, генерируя незашумленные и резкие изображения-модели, характер которых должен имитировать предполагаемый характер идеального изображения. Пусть, например, $F_I(x, y)$ описывает нерезкое и зашумленное изображение машинописной страницы; в этом случае должна использоваться модель в виде четкого и неискаженного шумами машинописного документа. Большой интерес представляет тот факт, что процедура слепой реставрации не требует явной оценки частотной характеристики нерезкости и оценки энергетического спектра шума. Эксперименты даютм ногообещающие результаты [9].

15.6. АНАЛИТИЧЕСКОЕ ПРОДОЛЖЕНИЕ СПЕКТРА

Инверсный фильтр, предназначенный для реставрации изображений, имеет частотную характеристику

$$\mathscr{H}_R(\omega_x, \omega_y) = 1/\mathscr{H}_D(\omega_x, \omega_y), \tag{15.6.1}$$

где $\mathscr{H}_D(\omega_x, \omega_y)$ — частотная характеристика изображающей системы с учетом пространственных искажений. В гл. 14 было показано, что осуществление такого фильтра часто оказывается затруднительным вследствие недопустимого увеличения шума в области высоких пространственных частот, где частотная характеристика искажений имеет малое значение. Более того, в любой реальной изображающей системе частотная характеристика $\mathscr{H}_D$ принимает нулевое значение вне некоторой области $(\omega_x^2 + \omega_y^2) > A$, где A — постоянная, поскольку такая система является дифракционно ограниченной. Поэтому создается впечатление, что инверсная фильтрация позволяет реставрировать составляющие изображения только в области пропускания изображающей системы, даже если шум отсутствует. Предположим, однако, что функция $F_I(x, y)$, описывающая идеальное изображение, ограничена по координатам x, y. Спектр $\mathscr{F}_I(\omega_x, \omega_y)$ должен быть неограниченным. Поскольку спектр идеального изображения является аналитической функцией, существует принципиальная возможность его однозначного продолжения за границы области пропускания изображающей системы, если он определен внутри этой области. Подвергая продолженный спектр обратному преобразованию Фурье, получают восстановленное изображение, соответствующее разрешению сверх дифракционного предела. Это и есть основная концепция «сверхразрешения» как метода повышения четкости изображения [12—15].

Один из возможных способов аналитического продолжения спектра изображения за дифракционный предел основан на применении сфероидальных волновых функций [16]. В одномерном случае эти функции обладают следующими свойствами. При заданных $T > 0$ и $\Omega > 0$ существуют счетное бесконечное множество действительных функций (собственных функций) $\Psi_i(t)$ и соответствующее множество действительных чисел (собственных значений) λ_i, такие, что

1) спектр Фурье функций $\Psi_i(t)$ тождественно равен нулю при $|\omega| > \Omega$;

2) функции $\Psi_i(t)$ ортонормальны на действительной оси:

$$\int_{-\infty}^{\infty} \Psi_i(t)\,\Psi_j(t)\,dt = \delta(i - j); \tag{15.6.2}$$

3) функции $\Psi_i(t)$ ортонормальны на интервале $-T/2 < t < T/2$:

$$\int_{-T/2}^{T/2} \Psi_i(t)\,\Psi_j(t)\,dt = \lambda_i \delta(i-j); \qquad (15.6.3)$$

4) для всех значений t

$$\lambda_i \Psi_i(t) = \int_{-T/2}^{T/2} [\sin\Omega(t-\alpha)/\pi(t-\alpha)]\,\Psi_i(\alpha)\,d\alpha. \qquad (15.6.4)$$

Сфероидальные волновые функции $\Psi_i(t)$ являются функциями произведения $c = T\Omega/2$, где T — временной интервал и Ω — гра-

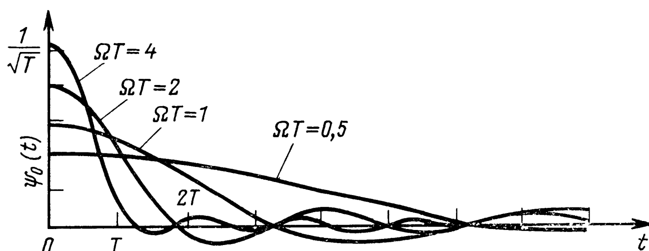


Рис. 15.6.1. Сфероидальные волновые функции.

ничная частота полосы пропускания. Рис. 15.6.1 иллюстрирует вид нескольких сфероидальных волновых функций. Предположим, что идеальное изображение имеет ограниченный размер, т. е.

$$F_I(x, y) = 0 \quad \text{при } |x| > L \quad \text{или} \quad |y| > L, \qquad (15.6.5)$$

где L — постоянная. Пусть $\mathscr{F}_I(\omega_x, \omega_y)$ — двумерный спектр Фурье функции $F_I(x, y)$, известный в некоторой области пропускания:

$$|\omega_x| \leqslant \omega_{xc}/2, \qquad (15.6.6а)$$

$$|\omega_y| \leqslant \omega_{yc}/2. \qquad (15.6.6б)$$

Разложив $\mathscr{F}_I(\omega_x, \omega_y)$ по системе спектральных сфероидальных волновых функций с переменными в виде пространственных частот, получим

$$\mathscr{F}_I(\omega_x, \omega_y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{ij} \Psi_i(\omega_x)\,\Psi_j(\omega_y) \qquad (15.6.7)$$

для всех ω_x, ω_y. Умножим обе части равенства (15.6.7) на $\Psi_m(\omega_x,)\Psi_n(\omega_y)$ и проинтегрируем результат в пределах об-

ласти пропускания. В результате получим выражение вида

$$\int_{-\omega_{xc}/2}^{\omega_{xc}/2}\int_{-\omega_{yc}/2}^{\omega_{yc}/2} \Psi_m(\omega_x)\Psi_n(\omega_y)\mathscr{F}_I(\omega_x, \omega_y)\,d\omega_x\,d\omega_y =$$

$$= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_{ij} \int_{-\omega_{xc}/2}^{\omega_{xc}/2}\int_{-\omega_{yc}/2}^{\omega_{yc}/2} \Psi_m(\omega_x)\Psi_n(\omega_y)\Psi_i(\omega_x)\Psi_j(\omega_y)\,d\omega_x\,d\omega_y. \tag{15.6.8}$$

Правая часть равенства (15.6.8) в силу (15.6.3) переходит в

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_{ij}\lambda_i\lambda_j\delta(i-m,\ j-m). \tag{15.6.9}$$

Следовательно,

$$a_{mn} = (1/\lambda_m\lambda_n)\int_{-\omega_{xc}/2}^{\omega_{xc}/2}\int_{-\omega_{yc}/2}^{\omega_{yc}/2} \Psi_m(\omega_x)\Psi_n(\omega_y)\mathscr{F}_I(\omega_x, \omega_y)\,d\omega_x\,d\omega_y. \tag{15.6.10}$$

Поскольку спектр $\mathscr{F}_I(\omega_x, \omega_y)$ в области пропускания определен, подстановка сфероидальных волновых функций и их собственных значений в (15.6.10) дает коэффициенты a_{mn}. Эти коэффициенты теперь можно подставить в формулу (15.6.7) и определить искомое разложение $\mathscr{F}_I(\omega_x, \omega_y)$ для всех ω_x, ω_y.

Метод аналитического продолжения спектра теоретически возможен, однако его трудно реализовать по следующим трем основным причинам:

1. Вычисление сфероидальных волновых функций и их собственных значений с высокой точностью, соответствующей достаточно большим произведениям протяженности этих функций на граничную частоту, сопряжено с большими трудностями.

2. Разложение (15.6.7), обеспечивающее аналитическое продолжение, удается вычислить лишь приближенно, т. е. с конечным числом членов, что связано с ошибками усечения.

3. Спектр, измеряемый в области пропускания изображающей системы и подставляемый в формулу (15.6.10) для нахождения коэффициентов разложения a_{mn}, фактически состоит из двух составляющих — спектра изображения и шумового спектра; таким образом, вычисление коэффициентов a_{mn} никогда не бывает абсолютно точным.

Главным препятствием на пути реализации процедуры аналитического продолжения всегда была и остается проблема шума. Расчеты показывают [15], что для аналитического продолжения

спектра за дифракционный предел с восстановлением всего нескольких пространственных частот требуется, чтобы отношение видеосигнал/шум было больше 1000 : 1. До сих пор было проведено лишь несколько машинных экспериментов с одномерными сигналами [13, 14].

15.7. РЕСТАВРАЦИЯ ИЗОБРАЖЕНИЙ С ПРОСТРАНСТВЕННО-ЗАВИСИМЫМИ ИСКАЖЕНИЯМИ

Пусть искажения, которые вносит изображащая система, можно промоделировать оператором, зависящим от линейного сдвига. В этом случае искаженное изображение описывается функцией

$$F_O(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_I(\alpha, \beta) H_D(x, y; \alpha, \beta) \, d\alpha \, d\beta, \qquad (15.7.1)$$

где $H_D(x, y; \alpha, \beta)$ — функция выходного изображения с координатами (x, y), создаваемого точечным источником с координатами (α, β). Основная задача реставрации формулируется следующим образом: при заданных $H_D(x, y; \alpha, \beta)$ и $F_O(x, y)$ найти оценку $F_I(x, y)$ посредством решения указанного интегрального уравнения. Общее решение этой задачи дает разложение $H_D(x, y; \alpha, \beta)$ по собственным функциям. Однако в большинстве ситуаций, возникающих в изображающих системах, собственные функции ядра либо неизвестны, либо просто не существуют.

Другой возможный подход к решению задачи — разбить изображение на небольшие фрагменты, в пределах которых искажение источника можно считать независимым от линейного сдвига, и воспользоваться уже рассмотренными методами. Общей теории выбора размеров фрагментов, при которых обеспечивается заданная точность результирующей оценки, пока не создано.

Наконец, существует особый класс задач реставрации изображений с искажениями, зависящими от линейного сдвига, в которых оператор, описывающий искажения, можно представить в виде совокупности трех операторов: первого оператора пространственных координатных искажений, оператора импульсного отклика, который не зависит от линейного сдвига, и второго оператора пространственных координатных искажений (рис. 15.7.1) [17]. В такой системе идеальное изображение сначала приобретает пространственные координатные искажения. В результате получается изображение, описываемое функцией

$$F_{O1}(x_1, y_1) = F_I[x(x_1\ y_1),\ y(x_1, y_1)]. \qquad (15.7.2)$$

Изображение на выходе системы с импульсным откликом, не зависящим от линейного сдвига, представляется функцией

$$F_{O2}(x_1, y_1) = F_{O1}(x_1, y_1) * H_D(x_1, y_1). \qquad (15.7.3)$$

Результирующее искаженное изображение описывается функцией

$$F_O(x_2, y_2) = F_{O2}[x_1(x_2, y_2), y_1(x_2, y_2)]. \qquad (15.7.4)$$

Если для обоих операторов пространственных координатных искажений можно найти однозначные обратные операторы, то

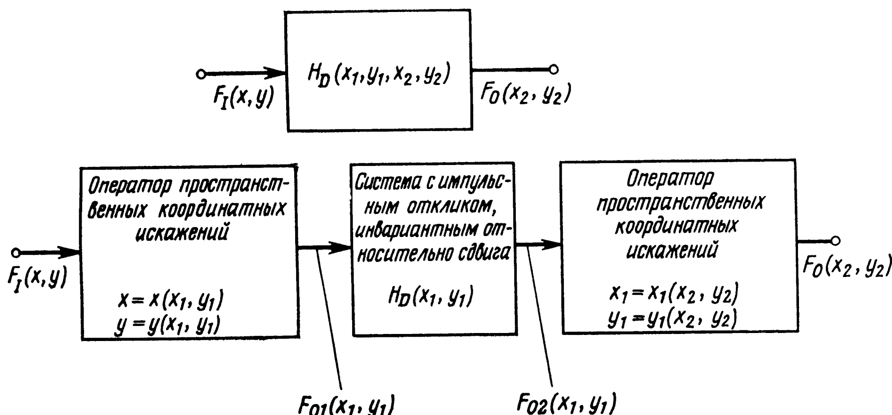


Рис. 15.7.1. Разложение оператора искажений изображающей системы с импульсным откликом, зависящим от линейного сдвига.

систему с искажением, зависящим от линейного сдвига, удается свести к системе с искажением, не зависящим от линейного сдвига; это позволяет получить решение, пользуясь уже рассмотренными методами. Указанный прием был с успехом применен для коррекции комы третьего порядка сферических линз [18, 19] и компенсации искажений изображения, вызванных движением [17].

15.8. РЕСТАВРАЦИЯ ИЗОБРАЖЕНИЙ МЕТОДОМ РЕКУРСИВНОЙ ФИЛЬТРАЦИИ

В последние несколько лет усиленно разрабатывались методы рекурсивной фильтрации, ориентированные на обработку сигналов в виде временны́х функций [20—22]. Недавно эти методы были приспособлены для реставрации изображений [23—25]. Решение задачи реставрации изображений на основе рекуррентного оценивания можно получить либо как результат обработки чисто двумерного сигнала, либо как результат обработки одномерного сигнала, возникающего при развертке.

15.8.1. ДВУМЕРНЫЙ ПОДХОД

Рассмотрим изображение, статистической характеристикой которого является автокорреляционная функция вида

$$R(\tau_1, \tau_2) = \sigma_s^2 \exp\{-\alpha_1 |\tau_1| - \alpha_2 |\tau_2|\}, \tag{15.8.1}$$

где τ_1 и τ_2 — приращения по горизонтали и вертикали соответственно. Случайное дискретное поле — массив элементов $F(j, k)$ — с такой автокорреляционной функцией можно сформировать, пользуясь рекуррентной формулой

$$F(j+1, k+1) = \rho_1 F(j+1, k) + \rho_2 F(j, k+1) - \\ - \rho_1\rho_2 F(j, k) + \sqrt{(1-\rho_1^2)(1-\rho_2^2)}\, u(j, k), \tag{15.8.2}$$

где ρ_1 и ρ_2 — коэффициенты корреляции между горизонтальными и вертикальными соседними элементами изображения соответственно, $u(j, k)$ — некоррелированное случайное поле, дисперсия которого равна дисперсии элементов изображения.

Если моделью наблюдаемого изображения, представленного массивом элементов $F_O(j, k)$, является сумма

$$F_O(j, k) = F_I(j, k) + N(j, k), \tag{15.8.3}$$

где $F_I(j, k)$ — элемент идеального изображения и $N(j, k)$ — элемент аддитивного шумового поля, то двумерная рекуррентная оценка идеального изображения [25] состоит из элементов

$$\hat{F}_I(j+1, k+1) = K_1(j, k)\hat{F}_I(j+1, k) + \\ + K_2(j, k)\hat{F}_I(j, k+1) + K_3(j, k)\hat{F}_I(j, k) + \\ + K_4(j, k) F_O(j+1, k+1), \tag{15.8.4}$$

где $K_i(j, k)$ — весовые коэффициенты. Формула (15.8.4) позволяет получить текущую оценку для каждого элемента изображения по оценкам для трех предыдущих элементов и текущему уровню яркости элемента наблюдаемого изображения с шумом. Оценка с минимальной среднеквадратической ошибкой получается, когда

$$K_1(j, k) = \rho_1, \tag{15.8.5а}$$

$$K_2(j, k) = \rho_2, \tag{15.8.5б}$$

$$K_3(j, k) = \rho_1\rho_2 + K_4(j, k) \tag{15.8.5в}$$

и $K_4(j, k)$ является сложной функцией ковариации ошибки оценивания [25]. Применение двумерной рекуррентной оценки (15.8.4) в целях реставрации ограничено случаем экспоненциальной автокорреляционной функции. Пока еще не разработан общий метод, позволяющий производить динамическое моделирование случайного поля с произвольной автокорреляционной функцией.

15.8.2. ОДНОМЕРНЫЙ ПОДХОД

В случае одномерной рекуррентной оценки наблюдаемый сигнал

$$f_O(t) = f_I(t) + n(t), \qquad (15.8.6)$$

получаемый в результате временнóй развертки изображения, моделируется суммой идеального «изображения» — функции $f_I(t)$ и аддитивного шума $n(t)$. Если автокорреляционная функция идеального изображения является экспоненциальной, то автокорреляционная функция наблюдаемого изображения может быть найдена [26]. При этом результирующая рекуррентная оценка j-го отсчета $f_I(t)$ определится как

$$\hat{f}_I(j) = k_1(j-1)\,\hat{f}_I(j) + k_2(j-1)\,f_O(j-1), \qquad (15.8.7)$$

где $k_1(j)$ и $k_2(j)$ — весовые коэффициенты, зависящие от автокорреляционной функции $f_I(t)$ [23, 24].

15.9. РЕСТАВРАЦИЯ ИЗОБРАЖЕНИЙ С ЦЕЛЬЮ ОСЛАБЛЕНИЯ ШУМА ЗЕРНИСТОСТИ ФОТОПЛЕНКИ

Шум зернистости фотопленки часто оказывается ограничивающим фактором при реставрации фотографических изображений, особенно при сильном увеличении. В разд. 13.6 представлен ряд моделей процесса фотографической записи, учитывающих шум зернистости фотопленки. Все эти модели обладают одним неприятным свойством: идеальное и наблюдаемое изображения связаны между собой нелинейным и нестационарным образом.

Уолкап и Чоенс [27] предложили использовать винеровскую фильтрацию для борьбы с шумом зернистости фотопленки в модели изображающей системы вида

$$F_O(x, y) = F_I(x, y) + \alpha\,[F_I(x, y)]^{1/3}\,N(x, y), \qquad (15.9.1)$$

где α — постоянная. Для этой модели была получена частотная характеристика реставрирующего фильтра, соответствующего случаю непрерывного изображения, в виде

$$\mathcal{H}_R(\omega_x, \omega_y) = \frac{\mathcal{W}_{F_I}(\omega_x, \omega_y)}{\mathcal{W}_{F_I}(\omega_x, \omega_y) + \alpha^2 E\left\{[F_I(x, y)]^{2/3}\right\}}, \qquad (15.9.2)$$

где $\mathcal{W}_{F_I}(\omega_x, \omega_y)$ — энергетический спектр идеального изображения, $E(\cdot)$ — обозначение математического ожидания. При выводе формулы (15.9.2) было принято неявное допущение о стационарности первого и второго моментов случайного поля, представляющего идеальное изображение. Надери и Савчук [28] разработали процедуру винеровского оценивания дискретных изображений на основе модели фотографической записи изображения на рис. 13.6.2.

Эта модель учитывает химические эффекты фотографического процесса, такие, как нелинейность характеристической кривой и пограничные эффекты, обусловленные диффузией проявителя, а также шум зернистости фотопленки. Преимущество рассматри-

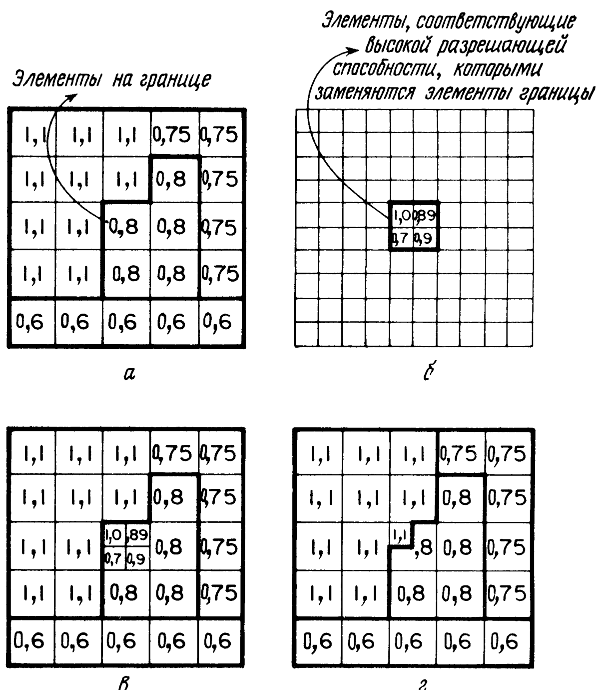


Рис. 15.9.1. Пример алгоритма сверхподавления шума, по Надери [31].

а — массив, соответствующий пониженной разрешающей способности; *б* — массив, соответствующий высокой разрешающей способности; *в* — замена элемента, принадлежащего границе; *г* — окончательный результат.

ваемой винеровской оценки состоит не только в том, что она основана на модели общего вида: благодаря присущей ей адаптивности она способна модифицироваться в соответствии с изменениями первого и второго моментов случайного поля, представляющего идеальное изображение.

Цвейг и др. [29, 30] разработали эвристический нелинейный метод реставрации малоконтрастных изображений с целью ослабления шума зернистости фотопленки (так называемое сверхподавление шума). При использовании этого метода входное изображение развертывается с высокой разрешающей способностью, а каждый его элемент квантуется с большим числом уровней. Затем получают изображение пониженной четкости, объединяя элементы

в непересекающиеся фрагменты размером 2×2. Конечно, четкое изображение имеет более резкие границы, чем изображение с пониженной четкостью, однако дисперсия шума последнего оказывается меньше. В случае белого шума дисперсия нечеткого изображения в четыре раза меньше, чем для четкого изображения, что является следствием пространственного усреднения элементов. Усредненное изображение повторно квантуется с использованием равномерной шкалы, причем шаг квантования выбирается равным учетверенному значению среднеквадратического отклонения шума. Благодаря такому выбору обеспечивается ошибка квантования 5% при гауссовом шуме. Полученные квантованные элементы нечеткого изображения исследуют в окрестности размером 3×3 элементов. Может оказаться, что все восемь периферийных элементов проквантованы с одним и тем же уровнем, а центральный элемент — с другим уровнем. В этом случае считают, что изолированный центральный элемент содержит ошибку, обусловленную шумом, и приписывают ему средний уровень периферийных элементов. Если центральный элемент нечеткого изображения лежит на границе (рис. 15.9.1), он разделяется на четыре элемента, соответствующие полной разрешающей способности; этим новым элементам приписываются уровни, зависящие как от уровней соответствующих исходных элементов четкого изображения, так и от уровней ближайших элементов нечетного изображения. Простой алгоритм заключается в том, что элементу, соответствующему высокой разрешающей способности, приписывают уровень одного из четырех связанных элементов (элементы «север» и «восток» или «север» и «запад» и т. д.), ближайший к уровню искомого элемента. Машинное моделирование алгоритма сверхподавления шума показало, что его можно с успехом использовать для реставрации малоконтрастных изображений, искаженных достаточно сильным шумом зернистости фотопленки [28]. Надери [31] обобщил идею алгоритма сверхподавления шума на случай реставрации высококонтрастных изображений, когда шум зернистости фотопленки зависит от содержания изображения.

Литература

1. Downing J. J., Data Sampling and Pulse Amplitude Molulation, in: Aerospace Telemetry, Stiltz H. L., Ed., Prentice-Hall, Englewood Cliffs, N. J., 1961.
2. Lesault R., Aliasing Problems in Two Dimensional Sampled Imagery, in: Perseption of Displayed Information, Biberman L., Ed., Plenum Press, New York, 1973. Chapter 7.
3. Andrews H. C., Patterson C. L., Digital Interpolation of Discrete Images, *IEEE Trans. Computers*, **C-25**, 2, 196—202 (1976).
4. O'Handley D. A., Green W. B., Recent Developments in Digital Image Processing at the Image Processing Laboratory at the Jet Propulsion Laboratory, *Proc. IEEE*, **60**, 7, 821—827 (1972). [Имеется перевод: О'Хендли, Грин,

Обзор достижений в области цифровой обработки изображений, проводимой лабораторией обработки изображения в Лаборатории реактивного движения. — ТИИЭР, 1972, т. 60, № 7, с. 84—92.]
5. Oppenheim A. V., Schafer R. W., Stockham T. G., Jr., Nonlinear Filtering of Multiplied and Convolved Signals, *Proc., IEEE*, **56**, 8, 1264—1291 (1968). [Имеется перевод: Оппенгейм, Шефер, Стокхэм, мл. Нелинейная фильтрация сигналов, представленных в виде произведения и свертки. — ТИИЭР, 1975, т. 56, № 8, с. 5—34.]
6. Andrews H. C., Hunt B. R., Digital Image Restoration, Prentice-Hall, Englewood Cliffs, N. J., 1976.
7. Papoulis A., Approximations of Point Spreads for Deconvolution, *J. Opt. Soc. Am.*, **62**, 1, 77—80 (1972).
8. Tatian B., Asymptotic Expansions for Correcting Truncation Error in Transfer-Function Calculation, *J. Opt. Soc. Am.*, **61**, 9, 1214—1224 (1971).
9. Stockham T. G., Jr., Cannon T. M., Ingebretsen P. B., Blind Deconvolution Through Digital Signal Processing, *Proc. IEEE*, **63**, 4, 678—692 (1975). [Имеется перевод: Стокхэм, мл., Кэннон, Ингебретсен. Цифровое восстановление сигналов посредством неопределенной инверсной свертки. — ТИИЭР, 1975, т. 63, № 4, с. 160—177.]
10. Cole E. R., The Removal of Unknown Image Blurs by Homomorphic Filtering, University of Utah, Computer Science Department, Report UTEC-CSc-74-029, June 1973.
11. Cannon T. M., Digital Image Deblurring by Nonlinear Homomorphic Filtering, University of Utah, Computer Science Department, Report UTEC-CSc-74-091, August 1974.
12. Frieden B. R., Band-unlimited Reconstruction of Optical Objects and Spectra, *J. Opt. Soc. Am.*, **57**, 8, 1013—1019. (1967).
13. Harris J. L., Image Evaluation and Restoration, *J. Opt. Soc. Am.*, **56**, 5, 569—574 (1966).
14. Huang T. S., Schreiber W. F., Tretiak O. J., Image Processing, *Proc. IEEE*, **59**, 11, 1586—1609 (1971). [Имеется перевод: Хуанг, Шрайбер, Третьяк. Обработка изображений. — ТИИЭР, 1971, т. 59, № 11, с. 59—89.]
15. Youla D. C., Restoration by the Method of Alternating Orthogonal Projections, Proceedings Symposium on Current Mathematical Problems in Image Science, Monterey, California, November 1976.
16. Slepian D., Pollak H. O., Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty—I, *Bell Syst. Tech. J.*, **40**, 43—63 (1961).
17. Sawchuk A. A., Space-Variant Image Motion Degradation and Restoration, *Proc. IEEE*, **60**, 7, 854—861 (1972). [Имеется перевод: Савчук. Пространственно-зависимые искажения изображения, вызванные движением, и реставрация изображений. — ТИИЭР, 1972, т. 60, № 7, с. 124—133.]
18. Robbins G. M., Huang T. S., Inverse Filtering for Linear Shift-Variant Imaging Systems, *Proc. IEEE*, **60**, 7, 862—872 (1972). [Имеется перевод: Роббинс, Хуанг. Принципы обратной фильтрации для линейных пространственно-зависимых систем отображения. — ТИИЭР, 1972, т. 60, № 7, с. 134—145.]
19. Sawchuk A. A., Spece-Variant Image Restoration by Coordinate Transformations, *J. Opt. Soc. Am.*, **64**, 2, 138—144 (1974).
20. Kalman R. E. A New Approach to Linear Filtering and Prediction Problems, *ASME Trans.* (*J. Basic Engineering*), **820**, 35—45 (March 1960).
21. Kalman R. E., Bucy R. C., New Results in Linear Filtering and Prediction Theory, *ASME Trans.* (*J. Basic Engineering*), **830**, 95—108 (March 1961).
22. Nahi N. E., Estimation Theory and Application, Wiley, New York, 1969.
23. Nahi N. E., Assefi T., Bayesian Recursive Image Estimation, in: Proceedings Two Dimensional Digital Signal Processing Conference, University of Missouri, October 1971.

24. Nahi N. E., Role of Recursive Estimation in Statistical Image Enhancement, *Proc. IEEE*, **60**, 7, 872—877 (1972). [Имеется перевод: Нахи. Применение метода рекуррентной оценки в статистической задаче улучшения качества изображения. — ТИИЭР, 1972, т. 60, № 7, с. 145—152.]
25. Habibi A., Two Dimensional Bayesian Estimate of Images, *Proc. IEEE*, **30**, 7, 878—883 (1972). [Имеется перевод: Хабиби. Двумерная бейесовская оценка изображений. — ТИИЭР, 1972, т. 60, № 7, с. 153—160.]
26. Franks L. E., A Model for Random Video Process, *Bell Syst. Tech. J.*, **45**, 609—630 (April 1966).
27. Walkup J. F., Choens R. C., Image Processing in Signal Dependent Noise, *Opt. Eng.*, **13**, 3, 258—266 (1974).
28. Naderi F., Sawchuk A. A., Nonlinear Detection and Estimation of Images Degraded by Film-Grain Noise, Proceedings SPIE/OSA Conference on Image Processing, Urbach J. C., Ed., Pacific Grove, California, February 1976, Vol. 74, pp. 17—24.
29. Zweig H. J., Barrett E. B., Hu P. C., Noise Cheating Image Enhancement, *J. Opt. Soc. Am.*, **65**, 11, 1347—1353 (1975).
30. Zweig H. J. et al., Experiment in Digital Restoration of Defocused Grainy Photographs by Noise Cheating and Fourier Techniques, Proceedings SPIE/OSA Conference on Image Processing, Urbac J. C., Ed., Pacific Grove, California, February 1976, Vol. 74, pp. 10—16.
31. Naderi F., Estimation and Detection of Images Degraded by Film-Grain Noise, University of Southern California, Image Processing Institute, Report USCIPI-690, September 1976.

Глава 16

КОРРЕКЦИЯ ЯРКОСТНЫХ, ЦВЕТОВЫХ И СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК ИЗОБРАЖАЮЩИХ СИСТЕМ

Реальные датчики изображения часто имеют неудовлетворительные спектральные или временны́е характеристики. Другой распространенный недостаток видеодатчиков и дисплеев — нелинейность характеристик передачи уровней. Подобные искажения удается в значительной степени скорректировать посредством обработки сигналов на выходе датчиков или на входе дисплеев. Реализация такой обработки обычно не представляет особых затруднений.

16.1. КОРРЕКЦИЯ ЯРКОСТНОЙ ПОЭЛЕМЕНТНОЙ НЕЛИНЕЙНОСТИ ВИДЕОДАТЧИКОВ

Существуют изображающие системы, в которых искажения можно представить в виде раздельных пространственной и поэлементной составляющих. В таких системах часто удается осуществить прямую компенсацию поэлементных искажений [1, 2]. Рассмотрим реальную изображающую систему, формирующую изображение, описываемое функцией $F_O(x, y, t)$, в соответствии с разделимой моделью:

$$F_O(x, y, t) = \mathcal{O}_P\{C, \mathcal{O}_D\{x, y, t, \lambda; C(x, y, t, \lambda)\}\}, \quad (16.1.1)$$

где $C(x, y, t, \lambda)$ — спектральная плотность энергии входного светового поля, $\mathcal{O}_P\{\cdot\}$ — поэлементная характеристика передачи уровней видеодатчика, $\mathcal{O}_D\{\cdot\}$ — оператор, определяемый пространственной, временно́й и спектральной характеристиками. Чтобы осуществить компенсацию яркостной нелинейности видеодатчика, наблюдаемое изображение следует пропустить через корректирующую систему, описываемую оператором реставрации $\mathcal{O}_R\{\cdot\}$, таким, что

$$\mathcal{O}_R\{\mathcal{O}_P\{\cdot\}\} = 1. \quad (16.1.2)$$

В случае оптических непрерывных изображений реализация нелинейного поэлементного реставрирующего оператора может оказаться затруднительной. Напротив, реставрация изображений, представленных в виде аналоговых электрических сигналов, выполняется весьма просто с помощью нелинейных усилителей;

легко осуществляется коррекция с помощью цифровых средств — обычных арифметических устройств или постоянных запоминающих устройств, реализующих таблицы соответствия.

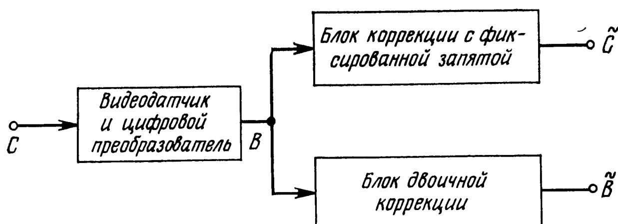


Рис. 16.1.1. Коррекция яркостной поэлементной нелинейности видеодатчика.

На рис. 16.1.1 представлена схема, иллюстрирующая методику коррекции яркостной поэлементной нелинейности. На входе видеодатчика находится точечный источник света с распределением C, которое преобразуется в двоичное число B, поступающее на вход вычислительной машины или цифровой системы связи. В некото-

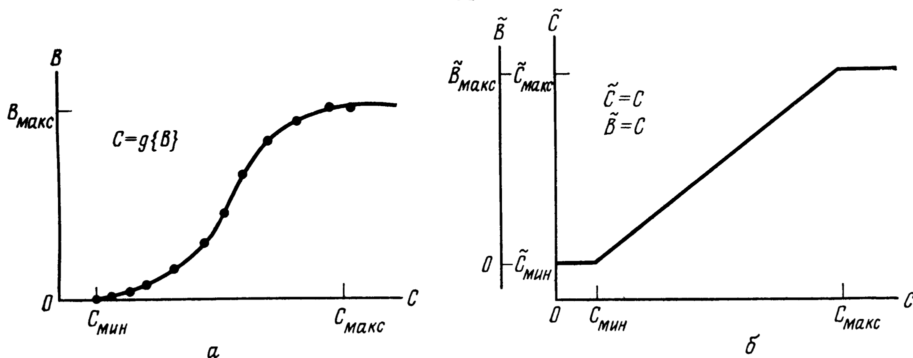


Рис. 16.1.2. Световые характеристики видеодатчика.
а — измеренная характеристика; *б* — скорректированная характеристика.

рых изображающих системах обработке подвергают непосредственно двоичные числа, в других — предпочитают предварительно формировать действительные машинные числа с фиксированной запятой, прямо пропорциональные яркости входного излучения. В первом случае используется блок двоичной коррекции, на выходе которого получается двоичное число $\tilde{B}$, прямо пропорциональное C. Во втором случае используется блок коррекции с фиксированной запятой, который дает число с фиксированной запятой $\tilde{C}$, равное C.

На рис. 16.1.2, *а* представлена типичная кривая зависимости измеренной величины B от яркости излучения на входе видеодат-

чика C; на рис. 16.1.2, *б* показана соответствующая скорректированная световая характеристика желательной формы. Исходную световую характеристику можно получить, пользуясь испытательной таблицей с полями известных яркостей; для каждого яркостного уровня определяется соответствующее двоичное число B. Для уменьшения ошибок, обусловленных шумом и измерительными погрешностями, следует многократно повторять измерения. Калибровка упростится, если двоичное представление яркости рассматривать как двоичное число с фиксированной запятой. Пусть, например, весь яркостный диапазон разбивается на 4096 уровней и кодируется с 12 разрядами. В этом случае будем пользоваться двоичным представлением вида

$$B = b_8 b_7 b_6 b_5 b_4 b_3 b_2 b_1 \cdot b_{-1} b_{-2} b_{-3} b_{-4}. \qquad (16.1.3)$$

Целая часть этого числа представляет уровни от 0 до 255, а дробная — доли каждого шага квантования, разбиваемого на 16 частей. При такой записи уровни выходного сигнала видеодатчика, выраженные в десятичной форме, будут находиться в диапазоне

$$255{,}9375 \leqslant B \leqslant 0{,}0. \qquad (16.1.4)$$

Определив ряд точек посредством измерения отдельных уровней яркости, строят по ним гладкую аналитическую кривую

$$C = g\{B\}. \qquad (16.1.5)$$

Искомые световые характеристики, выраженные через действительные и двоичные числа соответственно, имеют вид

$$\tilde{C} = C \qquad (16.1.6\text{а})$$

и

$$\tilde{B} = B_{\text{макс}} \frac{(C - C_{\text{мин}})}{(C_{\text{макс}} - C_{\text{мин}})}. \qquad (16.1.6\text{б})$$

Отсюда получаем выражения для скорректированных световых характеристик

$$\tilde{C} = g\{B\} \qquad (16.1.7\text{а})$$

и

$$\tilde{B} = B_{\text{макс}} \frac{(g\{B\} - C_{\text{мин}})}{(C_{\text{макс}} - C_{\text{мин}})}. \qquad (16.1.7\text{б})$$

Обычно пользуются нормированным диапазоном яркости от 0 до 1.

Точность калибровочной процедуры можно повысить, если сначала провести грубую калибровку, а затем повторить калибровку столько раз, сколько необходимо для нахождения уточненной скорректированной характеристики. Поскольку B — двоичное число, скорректированное значение яркости $\tilde{C}$ является действительным числом квантованной шкалы. Следует также заме-

тить, что скорректированное значение двоично-кодированной яркости B будет содержать ошибку округления, возникающую при вычислении правой части (16.1.7б). Таким образом, нелинейность кривой $C = g\{B\}$, построенной по точкам, и погрешность квантования, свойственная процессу преобразования в цифровую форму, могут привести к тому, что некоторые уровни скорректированной шкалы двоично-кодированной яркости не будут использоваться. Другими словами, гистограмма распределения уровней $\tilde{B}$, построенная для скорректированного изображения, может содержать пропуски. Чтобы минимизировать этот эффект, число уровней выходной шкалы должно быть меньше числа уровней входной шкалы. Например, при 12-разрядном кодировании B можно ограничиться только восемью разрядами кодирования $\tilde{B}$. Другое возможное решение — добавление к $\tilde{B}$ псевдослучайного шума, обеспечивающего более равномерное заполнение уровней квантования.

Многие видеодатчики с разверткой изображения, такие, как видиконные и диссекторные телекамеры, вносят яркостную поэлементную нелинейность, которая является пространственно-зависимой. В принципе коррекция яркостной поэлементной нелинейности может быть произведена для каждого элемента изображения в отдельности с использованием уже рассмотренных методов; при этом для каждого элемента должна быть известна откалиброванная световая характеристика. Однако технические трудности делают такую процедуру практически невозможной. Другой подход, часто дающий хорошие результаты, основан на моделировании пространственной неоднородности световой характеристики некоторой гладкой двумерной функцией $G(j, k)$ в пределах светочувствительной поверхности видеодатчика. При этом скорректированная пространственная световая характеристика $\tilde{F}(j, k)$ выражается в виде

$$\tilde{F}(j, k) = \frac{F(j, k)}{G(j, k)}, \qquad (16.1.8)$$

где $F(j, k)$ — исходная световая характеристика видеодатчика.

16.2. КОРРЕКЦИЯ ЯРКОСТНОЙ ПОЭЛЕМЕНТНОЙ НЕЛИНЕЙНОСТИ ДИСПЛЕЕВ

Яркостная поэлементная нелинейность дисплея в принципе корректируется так же, как и соответствующая нелинейность видеодатчика. Схема на рис. 16.2.1 иллюстрирует процедуру предыскажения величины яркости B двоично-кодированного изображения для формирования скорректированной величины яркости $\tilde{B}$, такой, что воспроизводимая яркость $\tilde{C}$ становится прямо

пропорциональной B. При этом предполагается, что процесс воспроизведения изображения может включать этап фотографиче-

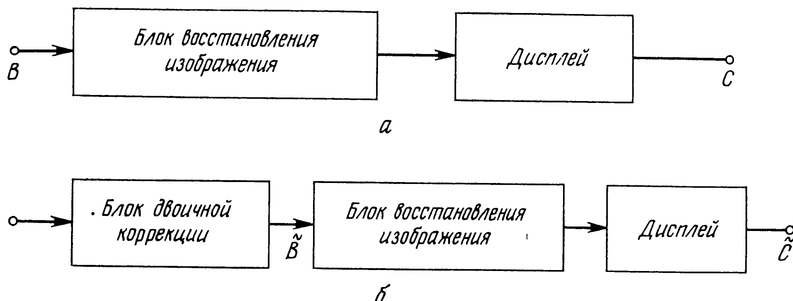


Рис. 16.2.1. Коррекция яркостной поэлементной нелинейности дисплея. *а* — без коррекции; *б* — с коррекцией.

ской записи светового поля. Желательная характеристика передачи уровней дисплея имеет вид

$$\tilde{C} = B\left[\frac{\tilde{C}_{\text{макс}} - \tilde{C}_{\text{мин}}}{B_{\text{макс}}}\right] + \tilde{C}_{\text{мин}}. \qquad (16.2.1)$$

Обычно максимальное и минимальное значения воспроизводимой яркости $\tilde{C}$ не выражаются в абсолютных единицах; чаще всего это нормированные коэффициенты пропускания или отражения. Измеренная характеристика передачи уровней дисплея и блока восстановления цифрового изображения моделируется нелинейной функцией

$$C = f\{B\}. \qquad (16.2.2)$$

Следовательно, линейная характеристика передачи уровней получается при

$$\tilde{B} = g\left\{B\left[\frac{\tilde{C}_{\text{макс}} - \tilde{C}_{\text{мин}}}{B_{\text{макс}}}\right] + \tilde{C}_{\text{мин}}\right\}, \qquad (16.2.3)$$

где $g(\cdot)$ — функция, обратная $f(\cdot)$.

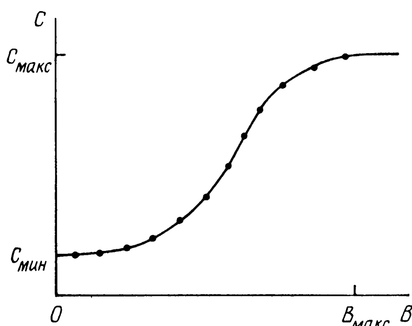


Рис. 16.2.2. Измеренная характеристика передачи уровней дисплея.

Процедуру экспериментального определения корректирующей функции $g(\cdot)$ поясним на примере изготовления фотографии с экрана дисплея. Прежде всего цифровым способом формируют изображение шкалы яркостей, охватывающей весь диапазон двоичных чисел. Обычно достаточно примерно 16 равномерно

распределенных уровней. Затем изготовляют фотоотпечаток и измеряют яркость отраженного света для каждого уровня шкалы. Результаты измерений наносят на график (рис. 16.2.2). По точкам проводят гладкую аналитическую кривую $B = g(C)$, определяю-

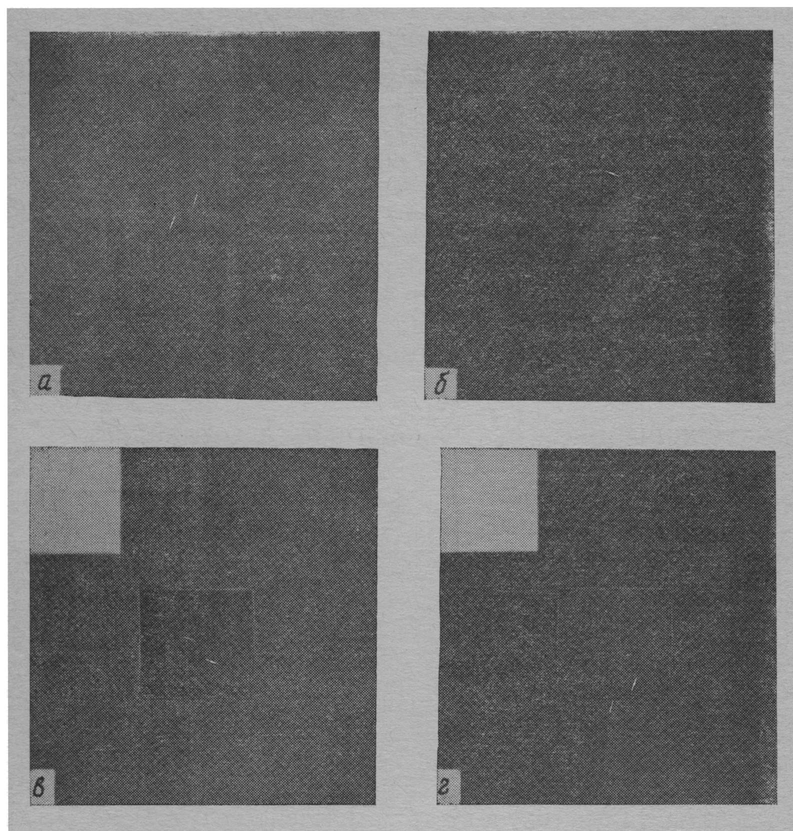


Рис. 16.2.3. Образцы изображений, подвергнутых пространственной яркостной коррекции.

а — некорректированное поле однородной яркости; *б* — скорректированное поле *а*; *в* — поле *а* с наклеенным вырезанным углом; *г* — поле *б* с наклеенным вырезанным углом.

щую требуемое преобразование. Важно, чтобы диапазон яркостей был проквантован на большое число уровней, поскольку это обеспечивает достаточную точность аппроксимации дискретного отображения $g(\cdot)$. Кроме того, число уровней должно быть достаточно большим, чтобы предотвратить появление ложных яркостных контуров вследствие неравномерного размещения уровней квантования в диапазоне яркости. В большинстве случаев доста-

точно пользоваться 8-разрядным представлением яркости B и 10-разрядным представлением яркости $\tilde{B}$.

Воспроизводящие устройства, подобные кинескопам, часто обладают пространственно-зависимой характеристикой передачи уровней. Как правило, яркость воспроизводимого изображения больше в центре и меньше по краям экрана. Пространственную неравномерность яркости изображения можно скомпенсировать, пользуясь, например, методом коррекции в соответствии с формулой (16.1.8). На рис. 16.2.3, *а* [3] представлена фотография (полученная с помощью видеодатчика бегущего луча), соответствующая постоянному уровню яркости. Яркость центральной части фотоснимка незначительно превышает яркость его периферийных областей. Различие в яркости более заметно на рис. 16.2.3, *в*, где приведен исходный фотоснимок (рис. 16.2.3, *а*), в центре которого наклеен вырезанный из него угол. Рис. 16.2.3, *б*, *г* иллюстрируют поля однородной яркости, полученные в результате пространственной яркостной коррекции.

16.3. ОЦЕНИВАНИЕ СПЕКТРАЛЬНОГО РАСПРЕДЕЛЕНИЯ ЭНЕРГИИ ИЗЛУЧЕНИЯ

При решении многих задач реставрации цветных и спектрозональных изображений должна быть найдена оценка спектрального распределения энергии излучения $C(\lambda)$ по серии наблюдаемых величин

$$x_i = \int C(\lambda) s_i(\lambda) d\lambda + n_i, \tag{16.3.1}$$

где $s_i(\lambda)$ — спектральная характеристика спектрозонального светофильтра, используемого для получения i-й величины, $i = 1, 2, 3 \dots P$. Член n_i представляет аддитивный шум или неопределенность измерения. Методы оценивания, описанные в гл. 14, позволяют найти решение и в рассматриваемом случае [4]. Первый шаг заключается в получении дискретного представления интеграла, что приводит к векторному уравнению

$$x_i = \mathbf{s}_i^T \mathbf{c} + n_i, \tag{16.3.2}$$

где $\mathbf{s}_i$ и $\mathbf{c}$ — векторы, составленные из Q отсчетов функций $s_i(\lambda)$ и $C(\lambda)$ соответственно. Совокупность P наблюдаемых изображений можно представить вектором размера $P \times 1$:

$$\mathbf{x} = \mathbf{S}\mathbf{c} + \mathbf{n}, \tag{16.3.3}$$

где вектор $\mathbf{s}_i^T$ находится в i-м столбце матрицы $\mathbf{S}$. Чтобы уменьшить погрешность численного интегрирования до разумных пределов, увеличивают число узловых точек. При этом система

уравнений (16.3.3) обычно оказывается существенно недоопределенной.

Оценку $\hat{\mathbf{c}}$ истинного спектрального распределения энергии $\mathbf{c}$ получают посредством обобщенного обращения:

$$\hat{\mathbf{c}} = \mathbf{S}^{-}\mathbf{x} = \mathbf{S}^{T}(\mathbf{S}\mathbf{S}^{T})^{-1}\mathbf{x}. \qquad (16.3.4)$$

Хотя применение метода обобщенного обращения матриц для нахождения оценки $\mathbf{c}$ гарантирует минимальную среднеквадратическую ошибку и минимальную норму, плохая обусловленность матрицы $\mathbf{S}$ и погрешности измерения могут привести к неустойчивым (осциллирующим) решениям. Обычно $\mathbf{c}$ имеет достаточно гладкий характер, поэтому при нахождении решения разумно воспользоваться теми или иными ограничениями на гладкость (см. разд. 14.7). Соответствующая оценка принимает вид

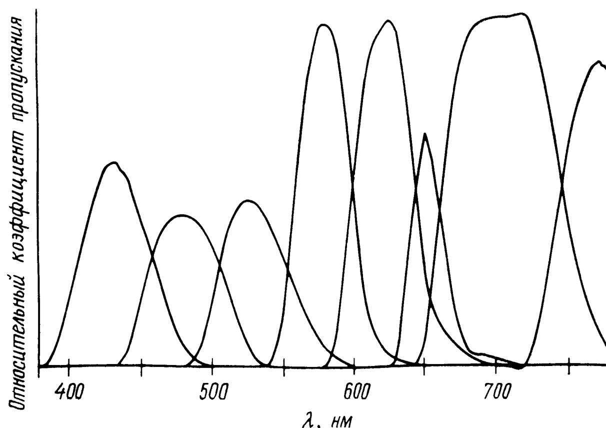


Рис. 16.3.1. Спектральные характеристики абсорбционных светофильтров.

$$\hat{\mathbf{c}} = \mathbf{M}^{-1}\mathbf{S}^{T}(\mathbf{S}\mathbf{M}^{-1}\mathbf{S}^{T})^{-1}\mathbf{x}, \qquad (16.3.5)$$

где $\mathbf{M}$ — матрица сглаживания, определенная соотношением (14.7.3). Третий возможный подход — применение методов винеровского оценивания, описанных в разд. 14.6. Винеровская оценка находится как

$$\hat{\mathbf{c}} = \mathbf{K}_{c}\mathbf{S}^{T}[\mathbf{S}\mathbf{K}_{c}\mathbf{S}^{T} + \mathbf{K}_{n}]^{-1}\mathbf{x}, \qquad (16.3.6)$$

где $\mathbf{K}_c$ — ковариационная матрица вектора $\mathbf{c}$, $\mathbf{K}_n$ — ковариационная матрица наблюдаемого аддитивного шума, независимого от $\mathbf{c}$. Поскольку $C(\lambda) \geqslant 0$ при любой длине волны, применение методов реставрации изображений с ограничениями (см. разд. 14.8) должно дать хорошие результаты.

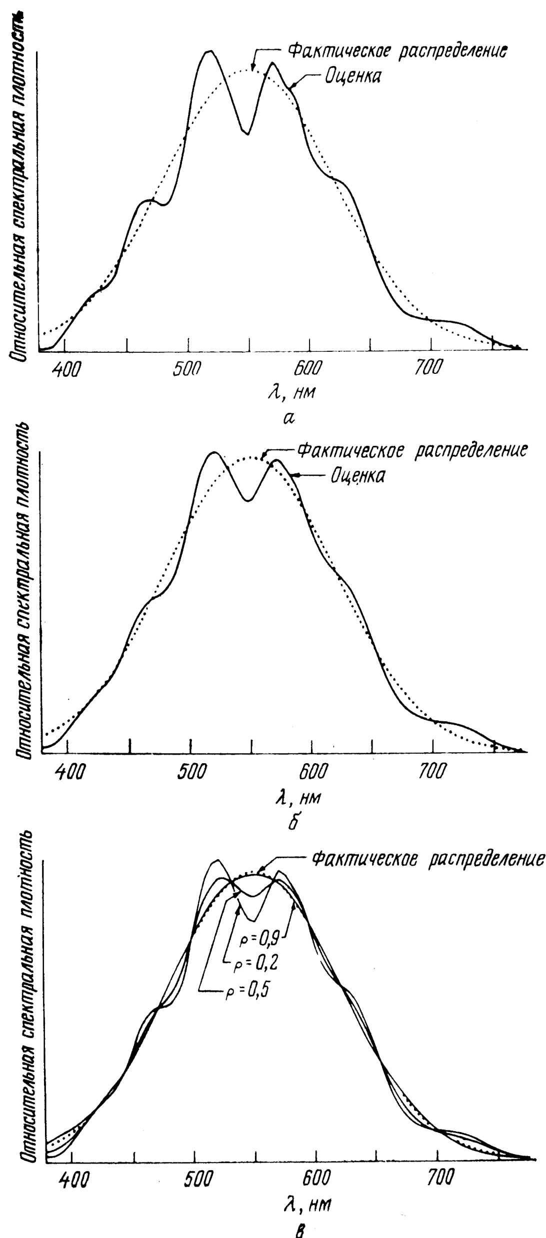


Рис. 16.3.2. Сравнение фактического спектрального распределения энергии излучения с его оценкой, полученной с помощью набора абсорбционных светофильтров (моделирование с помощью ЭВМ).

а — оценка методом псевдообращения матриц; *б* — оценка с использованием сглаживания; *в* — винеровская оценка (отношение сигнал/шум равно 1000).

Для сравнения указанных методов оценивания был проведен эксперимент, в котором с помощью ЭВМ моделировались измерения гауссова спектрального распределения энергии излучения с помощью набора различных светофильтров. На рис. 16.3.1 показаны спектральные характеристики этих светофильтров. Результаты измерений использовались в качестве спектрозональных наблюдаемых величин (16.3.1), по которым находилась оценка распределения $C(\lambda)$. Рис. 16.3.2 иллюстрирует эффективность трех методов оценивания: посредством псевдообращения матриц, с использованием сглаживания и винеровского. Винеровская оценка получена при допущениях, что шум белый, а распределение $C(\lambda)$ — реализация марковского процесса. Эксперименты показали, что минимальную среднеквадратическую ошибку аппроксимации исходного распределения $C(\lambda)$ оценкой $\hat{C}(\lambda)$ обеспечивает метод винеровского оценивания при коэффициенте корреляции марковского процесса $\rho = 0{,}9$ и отношении сигнал/шум, равном 1000.

В работе [4] рассмотрен пример спектральной калибровки датчика цветового видеосигнала методом оценивания спектрального распределения энергии излучения.

16.4. КОРРЕКЦИЯ СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК ВИДЕОДАТЧИКОВ И ДИСПЛЕЕВ

Как правило, спектральные характеристики датчиков изображающих систем определяются кривыми спектральной чувствительности используемых в них светочувствительных материалов, поэтому спектральные характеристики очень редко являются оптимальными. Однако неподходящие спектральные характеристики можно корректировать.

Пусть яркость, которая соответствует распределению энергии $C(\lambda)$, равна

$$X = \int C(\lambda)\, S(\lambda)\, d\lambda, \qquad (16.4.1)$$

где $S(\lambda)$ — спектральная характеристика видеодатчика. Требуемая яркость определяется как

$$Y = \int C(\lambda)\, V(\lambda)\, d\lambda, \qquad (16.4.2)$$

где $V(\lambda)$ — желательная спектральная характеристика. Операция коррекции спектральной характеристики сводится к отысканию некоторого функционального преобразования $f\{X\}$, которому

следует подвергнуть X, чтобы получить оценку $\hat{Y}$ искомой яркости Y. Таким образом, можно записать

$$\hat{Y} = f\{X\}. \qquad (16.4.3)$$

Корректирующее преобразование спектральных характеристик линейно только в том случае, если $S(\lambda)$ — линейная функция $V(\lambda)$. Обычно это условие не выполняется. Однако простота линейных преобразований часто побуждает искать оценку

$$\hat{Y}(\lambda) = a_1 S(\lambda) + a_2, \qquad (16.4.4)$$

обеспечивающую минимум той или иной функции ошибки, например среднеквадратической ошибки

$$\mathscr{E} = \int [Y(\lambda) - \hat{Y}(\lambda)]^2 d\lambda. \qquad (16.4.5)$$

В следующем разделе представлены результаты линейной коррекции спектральных характеристик видеодатчика. Хорошие результаты могла бы дать нелинейная коррекция спектральных характеристик, однако эффективные нелинейные методы до сих пор не разработаны.

Принципы коррекции спектральных характеристик видеодатчиков можно обобщить на случай коррекции соответствующих характеристик дисплеев. Рассмотрим некоторый дисплей со спектральной плотностью излучения $E_A(\lambda)$ и идеальный дисплей с желаемой спектральной плотностью излучения $E_I(\lambda)$. Спектральное распределение энергии для элемента изображения $F(j, k)$ определяется произведением $F(j, k)\, E_A(\lambda)$. Элемент исправленного изображения формируется в результате линейного преобразования вида

$$G(j, k) = a_1 F(j, k) + a_2, \qquad (16.4.6)$$

где a_1 и a_2 — постоянные, такие, что среднеквадратическое отклонение спектрального распределения энергии $G(j, k)\, E_A(\lambda)$ от распределения $F(j, k)\, E_I(\lambda)$ минимально для всех длин волн.

16.5. ОЦЕНИВАНИЕ КООРДИНАТ ЦВЕТА ВИДЕОДАТЧИКОВ

Цветной видеодатчик вырабатывает сигналы X_R, X_G, X_B, пропорциональные содержанию красного, зеленого и синего в передаваемой сцене:

$$X_R = \int C(\lambda)\, S_R(\lambda)\, d\lambda, \qquad (16.5.1а)$$

$$X_G = \int C(\lambda)\, S_G(\lambda)\, d\lambda, \qquad (16.5.1б)$$

$$X_B = \int C(\lambda)\, S_B(\lambda)\, d\lambda, \qquad (16.5.1в)$$

где $C(\lambda)$ — спектральное распределение энергии входного излучения, $S_R(\lambda)$, $S_G(\lambda)$, $S_B(\lambda)$ — спектральные характеристики красного, зеленого и синего каналов видеодатчика соответственно. Согласно принципам колориметрии, световое излучение, характеризуемое распределением $C(\lambda)$, можно воспроизвести при помощи набора основных цветов с фиксированными спектральными распределениями энергии. Количества основных цветов, обеспечивающие цветовое уравнивание, представляют собой координаты цвета

$$T_R = \int C(\lambda)\, T_{SR}(\lambda)\, d\lambda, \tag{16.5.2а}$$

$$T_G = \int C(\lambda)\, T_{SG}(\lambda)\, d\lambda, \tag{16.5.2б}$$

$$T_B = \int C(\lambda)\, T_{SB}(\lambda)\, d\lambda, \tag{16.5.2в}$$

где $T_{SR}(\lambda)$, $T_{SG}(\lambda)$, $T_{SB}(\lambda)$ — удельные координаты для соответствующих основных цветов. Сигналы видеодатчика X_R, X_G, X_B будут координатами цвета для некоторого набора основных цветов в том случае, если спектральные характеристики $S_R(\lambda)$, $S_G(\lambda)$, $S_B(\lambda)$ видеодатчика представляют собой линейную комбинацию удельных координат $T_{SR}(\lambda)$, $T_{SG}(\lambda)$, $T_{SB}(\lambda)$. Однако это условие выполняется в редких случаях. Рассмотрим теперь методы оценивания координат цвета на основе измеренных выходных сигналов цветного видеодатчика [5]. Простейший метод заключается в получении оценок удельных координат с помощью линейных операций

$$\hat{T}_{SR}(\lambda) = m_{11} S_R(\lambda) + m_{12} S_G(\lambda) + m_{13} S_B(\lambda), \tag{16.5.3а}$$

$$\hat{T}_{SG}(\lambda) = m_{21} S_R(\lambda) + m_{22} S_G(\lambda) + m_{23} S_B(\lambda), \tag{16.5.3б}$$

$$\hat{T}_{SB}(\lambda) = m_{31} S_R(\lambda) + m_{32} S_G(\lambda) + m_{33} S_B(\lambda), \tag{16.5.3в}$$

где $m_{i,j}$ — постоянные. Умножение обеих частей уравнений (16.5.3) на $C(\lambda)$ и интегрирование в пределах видимого спектра приводит к матричному уравнению

$$\begin{bmatrix} \hat{T}_R \\ \hat{T}_G \\ \hat{T}_B \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} X_R \\ X_G \\ X_B \end{bmatrix}. \tag{16.5.4}$$

В компактной записи это уравнение имеет вид

$$\hat{\mathbf{t}} = \mathbf{M}\mathbf{x}. \tag{16.5.5}$$

Таблица 16.5.1

Сравнение трех методов оценивания координат цвета по верности цветопередачи

Цвет	Фактическая яркость (отн.)	Аппроксимация удельных координат			Точное цветовое уравнивание по опорным цветам (*A*)			Точное цветовое уравнивание по опорным цветам (*B*)		
		яркость (отн)	$\mathscr{E}_Y$, %	$\mathscr{E}_C$, %	яркость (отн)	$\mathscr{E}_Y$, %	$\mathscr{E}_C$, %	яркость (отн)	$\mathscr{E}_Y$, %	$\mathscr{E}_C$, %
1. Телесный	0,331	0,329	0,48	7,36	0,349	5,55	5,81	0,331	0,08	4,86
2. Травы	0,121	0,129	6,28	9,01	0,128	5,59	3,34	0,123	0,05	0,18
3. Голубого неба	0,466	0,463	0,73	7,19	0,482	3,49	11,40	0,466	0,09	0,08
4. Красного кирпича	0,141	0,139	1,33	8,32	0,150	6,08	3,83	0,141	0,00	0,03
5. Фиолетового цветка	0,181	0,182	0,55	18,80	0,207	14,20	13,50	0,198	9,17	16,80
6. Синего красителя	0,087	0,081	6,99	29,80	0,087	0,30	0,15	0,086	1,11	21,08
7. Зеленого красителя	0,234	0,244	4,48	7,83	0,233	0,38	0,06	0,220	5,85	7,61
8. Желтого красителя	0,440	0,453	2,85	5,03	0,453	2,96	5,94	0,426	3,22	5,56
9. Красного красителя	0,105	0,086	18,10	5,97	0,105	0,13	0,02	0,099	5,49	4,53
10. Серый	0,500	0,500	0,01	8,86	0,517	3,38	3,58	0,494	1,24	2,12

$$\mathscr{E}_Y^2 = \left[\frac{Y(C) - \hat{Y}(C)}{Y(C)}\right]^2, \qquad \mathscr{E}_C = \left[\frac{u(C) - \hat{u}(C)}{u(C)}\right]^2 + \left[\frac{v(C) - \hat{v}(C)}{v(C)}\right]^2$$

Один возможный подход к определению оценок координат цвета $\hat{\mathbf{t}}$ по сигналам видеодатчика $\mathbf{x}$ сводится к выбору элементов матрицы преобразования $\mathbf{M}$ таким образом, чтобы минимизировать среднеквадратические ошибки для разностей $T_{SR}(\lambda) - \hat{T}_{SR}(\lambda)$,

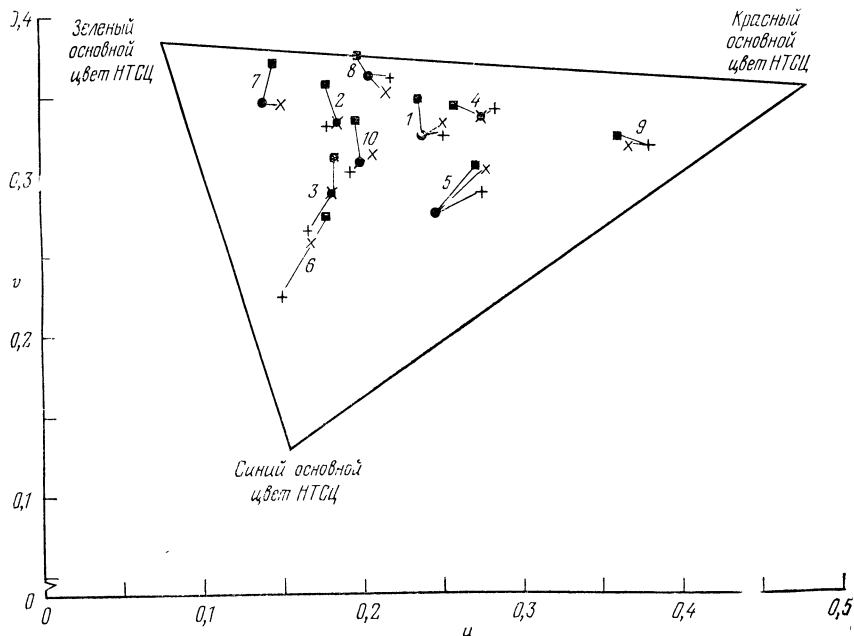


Рис. 16.5.1. Координаты цветности испытательных цветов, рассчитанные по оценкам координат цвета.

● фактические координаты цветности; ■ оценки координат цвета получены методом аппроксимации удельных координат в равноконтрастной колориметрической системе; + оценки координат цвета получены методом точного цветового уравнивания по трем цветам (6, 7 и 9); × оценки координат цвета получены методом точного цветового уравнивания по трем цветам (2, 3 и 4).

$T_{SG}(\lambda) - \hat{T}_{SG}(\lambda)$, $T_{SB}(\lambda) - \hat{T}_{SB}(\lambda)$ в пределах видимого спектра. Такая процедура была применена для оценивания цветопередачи десяти испытательных цветов, указанных в табл. 16.5.1 [6], при использовании эмульсионных слоев цветной фотопленки с типовыми характеристиками спектральной чувствительности. Координаты цветности этих цветов показаны на рис. 16.5.1 квадратами. Соответствующие относительные яркости представлены в табл. 16.5.1. Как видно, оцененные относительные яркости мало отличаются от фактических относительных яркостей для всех испытательных цветов; напротив, координаты цветности этих цветов, рассчитанные по оценкам координат цвета, заметно отличаются от фактических значений.

Второй подход к оцениванию координат цвета предполагает выбор трех опорных цветов и вычисление коэффициентов m_{ij} для точного колориметрического воспроизведения этих цветов. По-видимому в этом случае и любые другие цвета будут иметь координаты, существенно не отличающиеся от колориметрически правильных значений. Как правило, с наибольшими ошибками в цветности воспроизводятся высоконасыщенные цвета. Учитывая это, в качестве опорных из табл. 16.5.1 были выбраны красный (№ 9), зеленый (№ 7) и синий (№ 6) испытательные цвета. Для этих трех цветов $[C_i]$ с известными координатами цвета $U(C_i)$, $V(C_i)$, $W(C_i)$ были составлены матричные уравнения

$$\mathbf{t}(C_6) = \mathbf{Mx}(C_6), \qquad (16.5.6а)$$

$$\mathbf{t}(C_7) = \mathbf{Mx}(C_7), \qquad (16.5.6б)$$

$$\mathbf{t}(C_9) = \mathbf{Mx}(C_9). \qquad (16.5.6в)$$

Совместное решение этих уравнений дает значения коэффициентов m_{ij}. С помощью матрицы преобразования, составленной по уравнениям (16.5.6), были вычислены координаты цветности всех десяти испытательных цветов, показанные на рис. 16.5.1. Вычисленные относительные яркости представлены в табл. 16.5.1. Вся процедура была повторена для менее насыщенных опорных цветов — цвета красного кирпича (№ 4), цвета травы (№ 2) и цвета голубого неба (№ 3). Полученные результаты также представлены на рис. 16.5.1 и в табл. 16.5.1. Как и ожидалось, оценки координат цветности для опорных цветов оказались точными. Для обоих наборов трех опорных цветов оценки координат цветности других (не опорных) цветов отклонялись от истинных значений гораздо меньше, чем при оценивании удельных координат методом, описанным ранее. Очевидна необходимость компромисса при выборе опорных цветов. Действительно, если из данного множества воспроизводимых цветов взять максимально насыщенные цвета, то менее насыщенные неопорные цвета, как правило, будут воспроизводиться с повышенной насыщенностью. Если же в качестве опорных использовать цвета средней насыщенности, то будет наблюдаться снижение насыщенности предельно насыщенных цветов.

Третий подход к линейному оцениванию координат цвета основан на дискретных методах оценивания, рассмотренных в гл. 14. Прежде всего интегральные выражения (16.5.1) и (16.5.2) следует выразить в дискретном виде

$$\mathbf{x} = \mathbf{Sc} \qquad (16.5.7)$$

и

$$\mathbf{t} = \mathbf{T}_S\mathbf{c}, \qquad (16.5.8)$$

где $\mathbf{c}$ — вектор, составленный из отсчетов спектрального распределения энергии $C(\lambda)$, $\mathbf{S}$ и $\mathbf{T}_S$ — матрицы размера $3 \times Q$, строки

которых образованы отсчетами спектральных характеристик видеодатчика и ординатами кривых сложения соответственно. Требуется найти матрицу линейного преобразования **M**, минимизирующую среднеквадратическое отклонение истинных значений **t** от оценок $\hat{\mathbf{t}} = \mathbf{Mx}$. Искомую оценку можно записать в виде

$$\hat{\mathbf{t}} = \mathbf{T}_S \hat{\mathbf{c}}, \tag{16.5.9}$$

где $\hat{\mathbf{c}}$ — оценка спектрального распределения энергии, полученная по наблюдаемому сигналу видеодатчика **x**. Таким образом, задача оценивания координат цвета видеодатчиков эквивалентна фундаментальной задаче оценивания спектрального распределения энергии, которая рассматривалась в разд. 16.3.

16.6. ОЦЕНИВАНИЕ ЭКСПОЗИЦИЙ ЦВЕТНОЙ ФОТОПЛЕНКИ

Цветную фотопленку можно рассматривать как измерительное средство, для которого экспозиции эмульсионных слоев в красном, зеленом и синем представляют собой результаты измерений спектрального распределения энергии некоторого падающего светового излучения. Однако экспозиции не поддаются прямому наблюдению, поэтому для их определения требуется использовать те или иные методы оценивания [5].

Согласно формуле (13.3.15), экспозиции слоев описываются следующими выражениями:

$$X_R = d_R \int C(\lambda) L_R(\lambda)\, d\lambda, \tag{16.6.1а}$$

$$X_G = d_G \int C(\lambda) L_G(\lambda)\, d\lambda, \tag{16.6.1б}$$

$$X_B = d_B \int C(\lambda) L_B(\lambda)\, d\lambda, \tag{16.6.1в}$$

где $C(\lambda)$ — спектральное распределение энергии света, d_R, d_G, d_B — экспозиционные постоянные, $L_R(\lambda)$, $L_G(\lambda)$, $L_B(\lambda)$ — характеристики спектральной чувствительности эмульсионных слоев. Обработка экспонированной фотопленки позволяет изготовить фотографический отпечаток или транспорант (негатив или диапозитив). Ниже будут рассматриваться только транспоранты, однако полученные для них результаты можно без труда обобщить на случай фотоотпечатков. Согласно формулам (13.3.16) — (13.3.18), спектральный коэффициент пропускания определяется как

$$\tau_T(\lambda) = 10^{-\{cD_{NC}(\lambda) + mD_{NM}(\lambda) + yD_{NY}(\lambda)\}}, \tag{16.6.2}$$

где c, m, y — относительные количества голубого, пурпурного и желтого красителей в транспоранте, $D_{NC}(\lambda)$, $D_{NM}(\lambda)$, $D_{NY}(\lambda)$ — спектральные оптические плотности красителей.

Предположим теперь, что видеодатчик с установленным в нем готовым транспорантом вырабатывает следующие три видеосигнала:

$$V_{\hat{R}} = k_{\hat{R}} \int H_{\hat{R}}(\lambda)\, \tau_T(\lambda)\, d\lambda, \tag{16.6.3а}$$

$$V_G = k_G \int H_G(\lambda)\, \tau_T(\lambda)\, d\lambda, \tag{16.6.3б}$$

$$V_B = k_B \int H_B(\lambda)\, \tau_T(\lambda)\, d\lambda, \tag{16.6.3в}$$

где k_R, k_G, k_B — коэффициенты усиления соответствующих каналов видеодатчика, $H_R(\lambda)$, $H_G(\lambda)$, $H_B(\lambda)$ — эквивалентные спектральные коэффициенты пропускания цветоделительных светофильтров.

Задачу оценивания можно сформулировать как задачу нахождения оценок экспозиций $\hat{X}_R$, $\hat{X}_G$, $\hat{X}_B$ на основе наблюдаемых напряжений V_R, V_G, V_B и априорно известной модели фотографического процесса при заданном критерии ошибки. Оценивание экспозиций оказывается трудной задачей даже при использовании относительно простого критерия среднеквадратической ошибки; это объясняется нелинейностью используемых процессов и интегральной природой экспозиций, определяющих уровни выходных видеосигналов. Для упрощения можно считать, что оптические плотности красителей и экспозиции эмульсионных слоев фотопленки связаны линейными соотношениями

$$c = K_{FC} + \gamma_C \lg(X_{\hat{R}}), \tag{16.6.4а}$$

$$m = K_{FM} + \gamma_M \lg(X_G), \tag{16.6.4б}$$

$$y = K_{FY} + \gamma_Y \lg(X_B), \tag{16.6.4в}$$

где K_{FC}, K_{FM}, K_{FY} — постоянные коэффициенты, γ_C, γ_M, γ_Y — тангенсы угла наклона характеристических кривых. При этом выражения для выходных видеосигналов принимают вид

$$V_{\hat{R}} = k_{\hat{R}} \int K(\lambda)\, H_{\hat{R}}(\lambda)\, (X_{\hat{R}})^{\Psi_C(\lambda)}\, (X_G)^{\Psi_M(\lambda)}\, (X_B)^{\Psi_Y(\lambda)}\, d\lambda, \tag{16.6.5а}$$

$$V_G = k_G \int K(\lambda)\, H_G(\lambda)\, (X_{\hat{R}})^{\Psi_C(\lambda)}\, (X_G)^{\Psi_M(\lambda)}\, (X_B)^{\Psi_Y(\lambda)}\, d\lambda, \tag{16.6.5б}$$

$$V_B = k_B \int K(\lambda)\, H_B(\lambda)\, (X_{\hat{R}})^{\Psi_C(\lambda)}\, (X_G)^{\Psi_M(\lambda)}\, (X_B)^{\Psi_Y(\lambda)}\, d\lambda, \tag{16.6.5в}$$

где

$$K(\lambda) = 10^{\{-[K_{FC}D_{NC}(\lambda)+K_{FM}D_{NM}(\lambda)+K_{FY}D_{NY}(\lambda)]\}},$$

$$\Psi_C(\lambda) = \gamma_C D_{NC}(\lambda),$$

$$\Psi_M(\lambda) = \gamma_M D_{NM}(\lambda),$$

$$\Psi_Y(\lambda) = \gamma_Y D_{NY}(\lambda).$$

Формулировка задачи в виде уравнений (16.6.5) позволяет минимизировать ошибку оценивания различными численными методами, такими, как итерационный метод Ньютона—Рафсона, метод наискорейшего спуска и динамическое программирование. Чтобы получить достаточно хорошую начальную оценку, следует принять, что цветоделительные светофильтры имеют узкополосные спектральные характеристики [1]).

При этом уравнения для выходных сигналов видеодатчика принимают вид

$$V_{\hat{R}} = A_{\hat{R}}(\hat{X}_{\hat{R}})^{\Psi_C(\lambda_R)}(\hat{X}_G)^{\Psi_M(\lambda_R)}(\hat{X}_B)^{\Psi_Y(\lambda_R)}, \qquad (16.6.6а)$$

$$V_G = A_G(\hat{X}_{\hat{R}})^{\Psi_C(\lambda_G)}(\hat{X}_G)^{\Psi_M(\lambda_G)}(\hat{X}_B)^{\Psi_Y(\lambda_G)}, \qquad (16.6.6б)$$

$$V_B = A_B(\hat{X}_{\hat{R}})^{\Psi_C(\lambda_B)}(\hat{X}_G)^{\Psi_M(\lambda_B)}(\hat{X}_B)^{\Psi_Y(\lambda_B)}, \qquad (16.6.6в)$$

где

$$A_{\hat{R}} = k_{\hat{R}}K(\lambda_{\hat{R}})\int H_{\hat{R}}(\lambda)\,d\lambda,$$

$$A_G = k_G K(\lambda_G)\int H_G(\lambda)\,d\lambda,$$

$$A_B = k_B K(\lambda_B)\int H_B(\lambda)\,d\lambda.$$

Входящие в формулы (16.6.6) величины λ_R, λ_G, λ_B представляют собой длины волн, соответствующие центрам тяжести фигур, образованных кривыми $H_R(\lambda)$, $H_G(\lambda)$, $H_B(\lambda)$ светофильтров; символ $\wedge$ указывает на использование аппроксимации. Взятие натурального логарифма от обеих частей равенства (16.6.6) приводит к следующему матричному уравнению:

$$\begin{bmatrix} \ln V_{\hat{R}} \\ \ln V_G \\ \ln V_B \end{bmatrix} = \begin{bmatrix} \Psi_C(\lambda_{\hat{R}}) & \Psi_M(\lambda_{\hat{R}}) & \Psi_Y(\lambda_{\hat{R}}) \\ \Psi_C(\lambda_G) & \Psi_M(\lambda_G) & \Psi_Y(\lambda_G) \\ \Psi_C(\lambda_B) & \Psi_M(\lambda_B) & \Psi_Y(\lambda_B) \end{bmatrix} \begin{bmatrix} \ln \hat{X}_{\hat{R}} \\ \ln \hat{X}_G \\ \ln \hat{X}_B \end{bmatrix} + \begin{bmatrix} \ln A_{\hat{R}} \\ \ln A_G \\ \ln A_B \end{bmatrix}. \qquad (16.6.7)$$

[1]) Это условие действительно соблюдается в лазерных видеодатчиках.

Таблица 16.6.1

Оценки экспозиций для испытательных цветов

Цвет	Экспозиции			Выходные сигналы видеодатчика			Оценки для узкополосных спектральных характеристик				Оценки, полученные итерационным методом Ньютона—Рафсона			
							оценки экспозиций			ошибка оценивания $\mathscr{E}_X$, %	оценки экспозиций			ошибка оценивания $\mathscr{E}_X$, %
	X_R	X_G	X_B	V_R	V_G	V_B	$\hat{X}_R$	$\hat{X}_G$	$\hat{X}_B$		$\hat{X}_R$	$\hat{X}_G$	$\hat{X}_B$	
1. Телесный	50,0	27,9	20,4	0,450	0,301	0,245	36,3	34,0	27,3	49,1	50,5	28,0	20,4	0,61
2. Травы	8,8	12,8	6,5	0,062	0,073	0,055	6,4	16,5	11,0	81,5	8,9	12,8	6,5	0,60
3. Голубого неба	37,5	51,9	76,0	0,499	0,611	0,771	31,5	55,4	58,8	28,5	37,5	51,9	75,7	0,36
4. Красного кирпича	25,3	10,8	6,1	0,166	0,080	0,053	18,8	15,1	10,5	86,7	25,3	10,8	6,1	0,36
5. Фиолетового цветка	40,3	15,1	24,5	0,300	0,159	0,227	28,9	22,0	28,1	55,2	40,1	15,1	24,5	0,62
6. Синего красителя	6,1	10,4	22,3	0,043	0,072	0,153	4,2	16,5	23,9	67,2	6,1	10,4	22,3	0,41
7. Зеленого красителя	6,1	25,4	7,0	0,065	0,138	0,078	5,3	26,6	13,2	89,6	6,1	25,3	7,0	0,39
8. Желтого красителя	43,8	40,3	6,4	0,469	0,371	0,110	35,3	41,5	14,7	131,0	43,5	40,2	6,4	0,71
9. Красного красителя	30,1	4,6	8,3	0,143	0,038	0,052	20,9	8,6	11,3	97,2	30,1	4,6	8,3	0,19
10. Серый	50,0	50,0	50,0	0,585	0,587	0,593	38,2	53,4	48,0	24,9	49,8	50,0	50,0	0,38

Чтобы найти оценки экспозиций, достаточно обратить (16.6.7) и взять антилогарифм от обеих частей полученного равенства. В табл. 16.6.1 для десяти испытательных цветов приведены значения фактических экспозиций, выходных сигналов видеодатчика и оцененных экспозиций, которые помещены вместе с соответствующими среднеквадратическими ошибками в колонке «Оценки для узкополосных спектральных характеристик». Представленные данные получены при использовании типовой цветной фотопленки и цветного видеодатчика бегущего луча с цветоделительными светофильтрами «Раттен». Следует отметить, что даже грубая аппроксимация, сделанная при выводе соотношений (16.6.6), приводит к достаточно небольшой ошибке оценивания (по фотографическим критериям).

Итерационный метод Ньютона—Рафсона позволяет уточнить оценки, полученные для узкополосных светофильтров. При использовании этого метода начальные оценки экспозиций подставляют в формулы (16.6.6) для получения соответствующих значений $\tilde{V}_R$, $\tilde{V}_G$, $\tilde{V}_B$ выходных сигналов видеодатчика. Затем находят разности $V_R - \tilde{V}_R$, $V_G - \tilde{V}_G$, $V_B - \tilde{V}_B$ и по ним вычисляют корректирующие члены ΔX_R, ΔX_G, ΔX_B, добавляемые к узкополосным оценкам экспозиций:

$$\begin{bmatrix} \Delta X_R \\ \Delta X_G \\ \Delta X_B \end{bmatrix} = \begin{bmatrix} \frac{\partial V_R}{\partial X_R} & \frac{\partial V_R}{\partial X_G} & \frac{\partial V_R}{\partial X_B} \\ \frac{\partial V_G}{\partial X_R} & \frac{\partial V_G}{\partial X_G} & \frac{\partial V_G}{\partial V_B} \\ \frac{\partial V_B}{\partial X_R} & \frac{\partial V_B}{\partial X_G} & \frac{\partial V_B}{\partial X_B} \end{bmatrix}^{-1} \begin{bmatrix} V_R - \tilde{V}_R \\ V_G - \tilde{V}_G \\ V_B - \tilde{V}_B \end{bmatrix}. \quad (16.6.8)$$

Итерации повторяются до тех пор, пока не будет достигнута требуемая точность. Результаты оценивания с использованием итерационного метода Ньютона—Рафсона приведены в табл. 16.6.1. В рассмотренном примере всего пять итераций потребовалось для того, чтобы среднеквадратическую ошибку оценивания экспозиций сделать меньше 1 %. Можно ускорить сходимость, если в разложения (16.6.7) и (16.6.8) ввести члены высокого порядка [5].

16.7. КОМПЕНСАЦИЯ СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК ИЗЛУЧЕНИЯ ЦВЕТНЫХ ДИСПЛЕЕВ

В разд. 16.5 описаны методы обработки выходных сигналов цветного видеодатчика, целью которой является нахождение оценок координат цвета спектрального распределения энергии в любой точке цветного изображения. Полученные оценки можно использовать для управления трехцветным дисплеем (например,

использующим цветной кинескоп с теневой маской), на экране которого будет воспроизводиться изображение передаваемой сцены. Обратимся теперь к практической задаче фотографирования воспроизводимого изображения и изготовления фотоотпечатка с приемлемыми колориметрическими характеристиками.

Рассмотрим блок-схему фотографической записи на рис. 16.7.1. Блок коррекции, управляемый координатами цвета T_1, T_2, T_3, вырабатывает исправленные координаты цвета $\tilde{T}_1$, $\tilde{T}_2$, $\tilde{T}_3$, которые

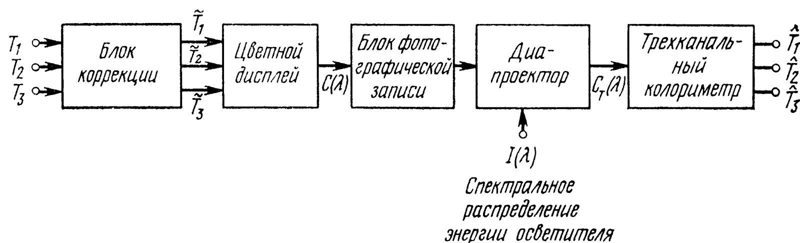


Рис. 16.7.1. Процесс изготовления и проецирования цветного транспоранта в системе воспроизведения цветных изображений.

в свою очередь управляют цветным дисплеем. Световое распределение, формируемое на экране последнего, записывается на трехслойную цветную фотопленку. Изготовленный цветной диапозитив проецируется на экран; координаты цвета $\hat{T}_1$, $\hat{T}_2$, $\hat{T}_3$ полученного изображения измеряются колориметром. Блок коррекции проектируется таким образом, чтобы свести к минимуму колориметрическое различие между цветами, определяемыми координатами T_1, T_2, T_3 и $\hat{T}_1$, $\hat{T}_2$, $\hat{T}_3$.

Спектральное распределение энергии излучения, создаваемое дисплеем:

$$C(\lambda) = \sum_{i=1}^{3} \tilde{T}_i P_{Si}(\lambda), \tag{16.7.1}$$

является взвешенной суммой спектральных распределений энергии излучения $P_{Si}(\lambda)$ основных цветов. Координаты цвета, соответствующие спектральным распределениям энергии проецируемого изображения, определяются как

$$\hat{T}_i = \int C_Г(\lambda) T_{Si}(\lambda) d\lambda, \tag{16.7.2}$$

где $T_{Si}(\lambda)$ — удельные координаты трехцветной системы, соответствующей координатам цвета T_i.

Известны два метода [7] коррекции спектральных характеристик излучения цветного дисплея: прямой и косвенный. Прямой метод предполагает разработку математической модели, основанной на уже рассмотренной модели цветного фотографического процесса (рис. 13.3.8); пользуясь такой моделью, можно шаг за шагом описать преобразование $C(\lambda)$ в $C_T(\lambda)$. Совокупность математической модели и уравнений (16.7.1) и (16.7.2) является детерминированным описанием взаимосвязи двух наборов координат цвета $\tilde{T}_i$ и $\hat{T}_i$. В идеальном случае эти наборы должны быть идентичными, однако на практике они таковыми не являются, поскольку характеристической кривой фотопленки присуща нелинейность, эмульсионные слои влияют друг на друга, а спектральные оптические плотности красителей перекрываются. Тем не менее количественное описание зависимости «вход—выход» достаточно для нахождения обратной зависимости. Прямой метод исследовался Уоллисом [7], который сумел существенно повысить колориметрическое качество цветной фотографической записи. Однако с прямым методом связаны серьезные вычислительные трудности, обусловленные относительной сложностью модели фотографического процесса. Косвенный метод коррекции спектральных характеристик излучения цветного дисплея сводится к нахождению некоторого матричного оператора **M**, минимизирующего различие между координатами цвета T_i и $\hat{T}_i$. Примером матричного оператора **M** может служить оператор с членами третьего порядка, осуществляющий преобразование вида

$$\begin{bmatrix} \tilde{T}_1 \\ \tilde{T}_2 \\ \tilde{T}_3 \end{bmatrix} = \mathbf{M} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_1^2 \\ T_2^2 \\ T_3^2 \\ T_1 T_2 \\ T_1 T_3 \\ T_2 T_3 \\ T_1^3 \\ T_2^3 \\ T_3^3 \end{bmatrix} + \begin{bmatrix} K_1 \\ K_2 \\ K_3 \end{bmatrix}, \qquad (16.7.3)$$

где K_i — постоянные. Усложняя преобразуемый вектор включением в него членов более высокого порядка, можно обеспечить

практически любую степень точности оценивания. Предсказание третьего порядка по формуле (16.7.3) оказалось достаточным для осуществления фотографической записи с экрана цветного кинескопа с маской [7]. Для того чтобы найти элементы корректирующей матрицы, необходимо определить координаты цвета большого числа цветов и соответствующие некорректированные координаты цвета $\hat{T}_i$ на выходе колориметра. Весовые элементы матрицы, найденные в результате ряда итераций, должны минимизировать выбранный функционал ошибки аппроксимации T_i оценкой $\hat{T}_i$.

ЛИТЕРАТУРА

1. Stockham T. G., Jr., A-D and D-A Converters: Their Effect on Digital Audio Fidelity, in: Digital Signal Processing, Rabiner L. R., Rader C. M., Eds., IEEE Press, New York, 1972, pp. 484—496.
2. O'Handley D. A., Green W. B., Recent Developments in Digital Image Processing at the Image Processing Laboratory at the Jet Propulsion Laboratory, *Proc. IEEE*, **60**, 7, 821—827 (1972). [Имеется перевод: О'Хендли, Грин. Обзор достижений в области цифровой обработки изображения, проводимой лабораторией обработки изображения в Лаборатории реактивного движения. — ТИИЭР, 1972, т. 60, № 7, с. 153—160.]
3. Frei W. (частное сообщение).
4. Pratt W. K., Mancill C. E., Spectral Estimation Techniques for the Spectral Calibration of a Color Image Scanner, *Appl. Opt.*, **14**, 11, 73—75 (1975).
5. Mancill C. E., Digital Color Image Restoration, University of Southern California, Image Processing Institute, USCIPI Report 630, August 1975.
6. Evans R. M., Hanson W. T., Jr., Brewer W. L., Principles of Color Photography, Wiley, New York, 1953.
7. Wallis R. H., Film Recording of Digital Color Images, University of Southern California, Image Processing Institute, USCIPI Report 570, May 1975.

Часть 5

Анализ изображений

Анализ изображения — выделение из изображения нужной информации с помощью автоматических или полуавтоматических приборов и систем. В литературе эта область имеет также другие названия: выделение данных из изображения, анализ сцен, описание изображений, автоматическое дешифрирование, распознавание изображений и т. д.

В отличие от других видов обработки изображений, таких, как кодирование, реставрация и улучшение качества изображений, результатом анализа изображений обычно является не картина, а ее числовое описание. Анализ изображений отличается также от классического распознавания образов тем, что системы анализа по определению не ограничиваются разделением областей сцены на фиксированное число классов, а предназначены для описания сложных сцен, разнообразие которых может быть настолько большим, что их нельзя описать с помощью заранее заданных терминов. Кроме того, в системах анализа изображений часто используются априорные сведения об изображаемых объектах и о их взаимных соотношениях. В системе анализа изображений могут также использоваться методы искусственного интеллекта для управления различными блоками системы и организации эффективного доступа к базе априорных сведений об объектах.

Глава 17

ВЫДЕЛЕНИЕ ПРИЗНАКОВ ИЗОБРАЖЕНИЯ

Признаком изображения называется его простейшая отличительная характеристика или свойство. Некоторые признаки являются естественными в том смысле, что они устанавливаются визуальным анализом изображения, тогда как другие, так называемые искусственные признаки, получаются в результате его специальной обработки или измерений. К естественным признакам относятся светлота (яркость) и текстура различных областей изображения, форма контуров объектов. Гистограммы распределения яркости и спектры пространственных частот дают примеры искусственных признаков.

17.1. ЯРКОСТНЫЕ ПРИЗНАКИ

Наиболее важным признаком изображения является светлота. Светлота выражается через такие величины, как яркость, спектральная интенсивность излучения, координаты цвета и т. д., которые будут называться яркостными признаками. Измерение яркостных признаков можно проводить либо в отдельных точках изображения, либо в их окрестностях. Например, средняя яркость окрестности точки (j, k) изображения размером $(2W+1) \times \times (2W+1)$ элементов определяется как

$$\overline{Y}(j, k) = [1/(2W+1)^2] \sum_{m=-W}^{W} \sum_{n=-W}^{W} Y(j+m, k+n). \quad (17.1.1)$$

Существует множество различных способов определения яркостных признаков. Можно использовать значения яркости или координат цвета непосредственно либо перейти к новым яркостным признакам, выполнив некоторое линейное, нелинейное или, возможно, необратимое преобразование.

Измерение яркостных признаков приобретает особую важность при выделении изображенных объектов (символическое описание) и при маркировке таких объектов (интерпретация). Эти вопросы рассматриваются в следующих главах.

17.2. ГИСТОГРАММНЫЕ ПРИЗНАКИ

В гл. 5 введено представление дискретного изображения — массива чисел $F(j, k)$ — как реализацию двумерного случайного процесса, описываемого совместными распределениями вероятностей. Разработаны методы оценки распределений вероятностей значений яркости изображения. Получающиеся при этом гистограммы можно использовать также для создания некоторого класса признаков изображения.

Распределение вероятностей значений яркости первого порядка можно определить как

$$P(b) \equiv \Pr\{F(j, k) = b\}, \qquad (17.2.1)$$

где $0 \leqslant b \leqslant L - 1$ — уровни квантования. Распределение частот первого порядка, оценивающее $P(b)$, описывается простым выражением

$$P(b) \approx N(b)/M, \qquad (17.2.2)$$

где M — полное число элементов изображения в окне с центром (j, k), а $N(b)$ — число элементов в окне, имеющих уровень b. Часто, исходя из предположения о стационарности, размер окна берется равным размеру всего изображения.

Форма гистограммы распределения частот дает много сведений о свойствах изображения. Например, узкая гистограмма указывает на низкий контраст, а бимодальная гистограмма предполагает наличие областей разной яркости. В работе [1] введены следующие характеристики, описывающие форму гистограмм первого порядка:

Среднее

$$\bar{b} = \sum_{b=0}^{L-1} bP(b). \qquad (17.2.3)$$

Дисперсия

$$\sigma_b^2 = \sum_{b=0}^{L-1} (b - \bar{b})^2 P(b). \qquad (17.2.4)$$

Коэффициент асимметрии

$$b_S = (1/\sigma_b^3) \sum_{b=0}^{L-1} (b - \bar{b})^3 P(b). \qquad (17.2.5)$$

Коэффициент эксцесса

$$b_K = (1/\sigma_b^4) \sum_{b=0}^{L-1} (b - \bar{b})^4 P(b) - 3. \qquad (17.2.6)$$

Энергия

$$b_N = \sum_{b=0}^{L-1} [P(b)]^2. \qquad (17.2.7)$$

Энтропия

$$b_E = -\sum_{b=0}^{L-1} P(b)\log_2[P(b)]. \qquad (17.2.8)$$

Слагаемое —3, введенное в выражение для коэффициента эксцесса, нормализует величину b_K так, что она равняется нулю для квантованной гауссовой плотности вероятности при стремящемся к нулю шаге квантования.

Гистограммные признаки второго порядка основаны на определении совместного распределения вероятностей пар элементов изображения. Рассмотрим два элемента $F(j, k)$ и $F(m, n)$ с координатами соответственно (j, k) и (m, n), причем вектор расстояния между ними определяется модулем r и углом θ относительно горизонтальной оси. Тогда совместное распределение значений яркости имеет вид

$$P(a, b) \equiv P_R\{F(j, k) = a,\ F(m, n) = b\}, \qquad (17.2.9)$$

где a и b — квантованные значения яркости. Вследствие дискретного представления изображения на прямоугольном растре параметры (r, θ) могут принимать лишь определенные дискретные значения. Распределение частот, оценивающее распределение второго порядка, определяется формулой

$$P(a, b) = N(a, b)/M, \qquad (17.2.10)$$

где M — полное число элементов в окне, а $N(a, b)$ — число случаев, когда $F(j, k) = a$ и $F(m, n) = b$.

Если пары элементов изображения сильно взаимосвязаны, то ненулевые элементы матрицы $\mathbf{P}(a, b)$ будут концентрироваться вдоль главной диагонали. Ниже перечислены характеристики этого процесса [1, 2]:

Автокорреляция

$$B_A = \sum_{a=0}^{L-1}\sum_{b=0}^{L-1} abP(a, b). \qquad (17.2.11)$$

Ковариация

$$B_C = \sum_{a=0}^{L-1}\sum_{b=0}^{L-1} (a-\bar{a})(b-\bar{b})P(a, b). \qquad (17.2.12)$$

Момент инерции

$$B_I = \sum_{a=0}^{L-1}\sum_{b=0}^{L-1} (a-b)^2 P(a, b). \qquad (17.2.13)$$

Средняя абсолютная разность

$$B_V = \sum_{a=0}^{L-1}\sum_{b=0}^{L-1} |a-b|\, P(a, b). \qquad (17.2.14)$$

«Обратная разность»

$$B_D = \sum_{a=0}^{L-1} \sum_{b=0}^{L-1} P(a, b)/[1 + (a-b)^2]. \qquad (17.2.15)$$

Энергия

$$B_N = \sum_{a=0}^{L-1} \sum_{b=0}^{L-1} [P(a, b)]^2. \qquad (17.2.16)$$

Энтропия

$$B_E = -\sum_{a=0}^{L-1} \sum_{b=0}^{L-1} P(a, b) \log_2 [P(a, b)]. \qquad (17.2.17)$$

Применение этих характеристик для текстурного анализа рассматривается в разд. 17.8.

17.3. ПРОСТРАНСТВЕННО-СПЕКТРАЛЬНЫЕ ПРИЗНАКИ

Спектральные коэффициенты, найденные в результате двумерного преобразования, определяют веса базисных изображений (двумерных базисных функций), соответствующих этому преобразованию, при которых взвешенная сумма базисных функций идентична изображению. Можно считать, что эти коэффициенты показывают степень корреляции соответствующих базисных функций с изображением. Если базисное изображение имеет ту же пространственную форму, что и признак, который необходимо обнаружить на изображении, то обнаружение признака можно выполнить просто путем наблюдения значения соответствующего спектрального коэффициента. Практическая сложность состоит в том, что объекты, которые необходимо обнаружить, часто имеют сложные форму и распределение яркостей и, следовательно, не соответствуют точно более простым яркостным образам, которые представляются базисными функциями большинства преобразований.

Лендерис и Стенли [3, 4] исследовали применение непрерывного двумерного преобразования Фурье, полученного с помощью когерентно-оптического устройства, для выделения признаков изображения. Оптическая система создает электрическое поле, пропорциональное спектру

$$\mathcal{F}(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(x, y) \exp\{-i(\omega_x x + \omega_y y)\}\, dx\, dy, \qquad (17.3.1)$$

где (ω_x, ω_y) — пространственные частоты. Оптическое считывающее устройство дает на выходе функции

$$\mathcal{M}(\omega_x, \omega_y) = |\mathcal{F}(\omega_x, \omega_y)|^2, \qquad (17.3.2)$$

значения которых пропорциональны интенсивности спектра. Следует отметить, что функции $\mathcal{F}(\omega_x, \omega_y)$ и $F(x, y)$ представляют собой пару, связанную однозначным преобразованием, тогда как функция $\mathcal{M}(\omega_x, \omega_y)$ неоднозначно связана с $F(x, y)$. Например, функция $\mathcal{M}(\omega_x, \omega_y)$ не изменяется, если начало координат на плоскости исходного изображения сдвигается. Для некоторых

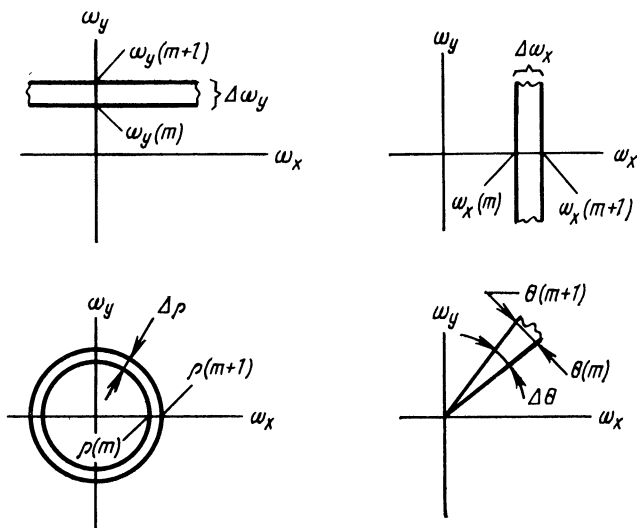


Рис. 17.3.1. Маски спектральных признаков.

применений инвариантность функции $\mathcal{M}(\omega_x, \omega_y)$ относительно сдвига может оказаться достоинством. Интегрирование функции $\mathcal{M}(\omega_x, \omega_y)$ по углу на плоскости пространственных частот дает пространственно-частотный признак, инвариантный относительно сдвига и вращения. Представив функцию $\mathcal{M}(\omega_x, \omega_y)$ в полярных координатах, получим этот признак в следующем виде:

$$\mathcal{N}(\rho) = \int_0^{2\pi} \mathcal{M}(\rho, \theta)\, d\theta, \tag{17.3.3}$$

где $\theta = \operatorname{arctg}(\omega_y/\omega_x)$ и $\rho^2 = \omega_x^2 + \omega_y^2$. Инвариантностью относительно изменения масштаба обладает признак

$$\mathcal{P}(\theta) = \int_0^{\infty} \mathcal{M}(\rho, \theta)\, \rho\, d\rho. \tag{17.3.4}$$

Если входное изображение имеет ограниченные размеры, то поле $\mathcal{F}(\omega_x, \omega_y)$, полученное в результате преобразования Фурье, будет

затухать в пределах определенного диапазона. Можно легко показать, что если функцию $F(x, y)$ умножить на функцию окна $W(x, y)$, которая равна единице внутри некоторого прямоугольника и нулю вне его, то фурье-спектр этого произведения равен свертке $\mathcal{F}(\omega_x, \omega_y)$ с фурье-спектром функции окна $\mathcal{W}(\omega_x, \omega_y)$. Затухание, вызванное ограниченными размерами окна, следует учитывать при определении относительных значений коэффициентов Фурье на различных пространственных частотах.

Для выделения признаков изображения яркостный образ $\mathcal{M}(\omega_x, \omega_y)$ обычно рассматривается в областях специфической формы. В качестве примера на рис. 17.3.1 определены области для следующих признаков Фурье:

Горизонтальная щель

$$S_1(m) = \int\limits_{\omega_y(m)}^{\omega_y(m+1)} \mathcal{M}(\omega_x, \omega_y)\, d\omega_y. \tag{17.3.5}$$

Вертикальная щель

$$S_2(m) = \int\limits_{\omega_x(m)}^{\omega_x(m+1)} \mathcal{M}(\omega_x, \omega_y)\, d\omega_x. \tag{17.3.6}$$

Кольцо

$$S_3(m) = \int\limits_{\rho(m)}^{(m+1)} \mathcal{M}(\rho, \theta)\, d\rho. \tag{17.3.7}$$

Сектор

$$S_4(m) = \int\limits_{\theta(m)}^{\theta(m+1)} \mathcal{M}(\rho, \theta)\, d\theta. \tag{17.3.8}$$

Для дискретного изображения, описываемого массивом чисел $F(j, k)$, в качестве источника признаков можно рассматривать непосредственно дискретный спектр Фурье

$$\mathcal{F}(u, v) = (1/N) \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} F(j, k) \exp\left\{\frac{-2\pi i}{N}(ux + vy)\right\} \tag{17.3.9}$$

при $u, v = 0, \ldots, N-1$. В этом случае признаки для горизонтальной щели, вертикальной щели, кольца и сектора можно определить аналогично выражениям (17.3.5)—(17.3.8). Описанную идею можно распространить также и на другие унитарные преобразования, такие, как преобразование Адамара и преобразование Хаара. На рис. 17.3.2 представлены распределения значений логарифма модуля спектра Фурье для нескольких геометрических форм.

Выделение признаков, представленных в виде спектральных коэффициентов, было исследовано в различных практических

задачах, в которых эти признаки использовались в качестве входных данных для системы распознавания образов. Спектральные

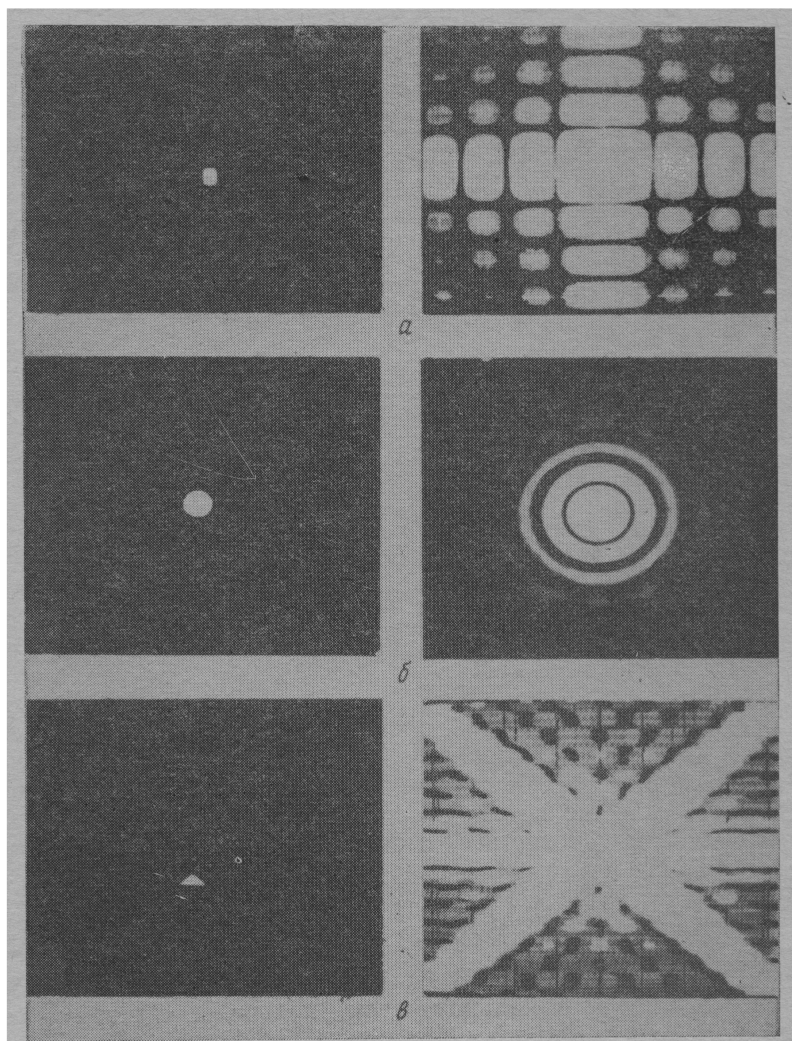


Рис. 17.3.2. Примеры дискретного спектра Фурье различных объектов. Показано распределение значений логарифма модуля спектра Фурье для $N = 256$.
а — прямоугольник; *б* — эллипс; *в* — треугольник.

признаки находят широкое применение — от классификации земельных ресурсов [5] до диагностики болезней по рентгеновским снимкам [6—8].

17.4. КОНТУРНЫЕ ПРИЗНАКИ

Резкие изменения (разрывы) яркости, координат цвета или параметров, характеризующих текстуру, являются важными простейшими признаками, поскольку они часто определяют очертания изображенных объектов. Локальные разрывы значений яркости называются яркостными перепадами, или яркостными контурами (luminance edge). Протяженные разрывы, называемые отрезками границы объекта, рассматриваются в гл. 18. В этом разделе изучаются яркостные перепады, разделяющие области

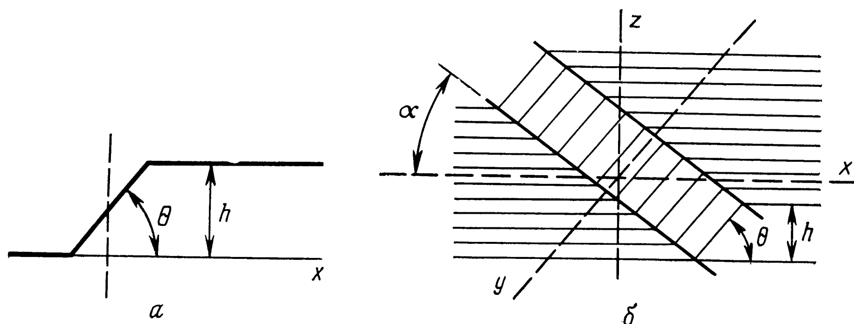


Рис. 17.4.1. Одномерный (*а*) и двумерный (*б*) перепады.

с почти одинаковой яркостью. Способы обнаружения границы между областями с разной текстурой рассмотрены в разд. 18.7. Обнаружение перепадов на цветных изображениях описано в разд. 17.6.

На рис. 17.4.1 показаны участки изображения с быстрым изменением яркости от низкого уровня к высокому в одномерном и двумерном случаях. В одномерном случае перепад характеризуется высотой, углом наклона и координатой центра склона. Перепад существует, если его угол наклона и высота больше некоторого заданного порога. Для двумерного случая важна также ориентация перепада по отношению к оси x. Идеальный детектор перепада при обработке областей изображения, представленных на рис. 17.4.1, должен указывать на наличие перепада в единственной точке, расположенной в центре склона.

Общий подход к обнаружению перепадов на одноцветном изображении иллюстрируется в виде блок-схемы на рис. 17.4.2. Исходное изображение, представленное массивом чисел $F(j, k)$, подвергается линейной или нелинейной обработке с тем, чтобы усилить перепады яркости. В результате образуется массив чисел $G(j, k)$, описывающий изображение с подчеркнутыми изменениями яркостей. Затем выполняется операция сравнения с порогом и

определяется положение элементов изображения с ярко выраженными перепадами. Если

$$G(j, k) < T_L(j, k), \qquad (17.4.1а)$$

то имеет место нисходящий перепад, а при

$$G(j, k) \geq T_U(j, k) \qquad (17.4.1б)$$

— восходящий перепад. Величины $T_L(j, k)$ и $T_U(j, k)$ представляют собой нижнее и верхнее пороговые значения. Эти значения можно сделать переменными в плоскости изображения для компенсации влияния сильных изменений яркости на результаты обнаружения перепадов.

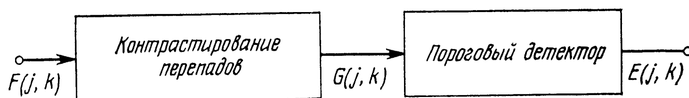


Рис. 17.4.2. Пороговая система обнаружения перепадов.

Выбор порога является одним из ключевых вопросов выделения перепадов. При слишком высоком уровне порога не будут обнаружены структурные элементы с низким контрастом. Наоборот, слишком низкий уровень порога явится причиной того, что шум будет ложно принят за перепад. Для обозначения положения перепадов на изображении часто формируют контурный препарат — массив элементов $E(j, k)$. Например, положение точек восходящих препаратов можно было бы отмечать белыми точками на черном фоне. Можно также точки восходящих препаратов отмечать белым цветом, нисходящих — черным, а остальные элементы изображения — некоторым средним уровнем яркости.

Другой важный подход к обнаружению перепада состоит в аппроксимации фрагмента реального изображения некоторым идеальным одно- или двумерным перепадом, показанным на рис. 17.4.1. Если аппроксимация оказывается достаточно точной, то считается, что перепад существует и ему приписываются параметры идеального перепада.

17.4.1. ЛИНЕЙНЫЕ МЕТОДЫ КОНТРАСТИРОВАНИЯ

До выполнения операции порогового обнаружения можно «подчеркнуть» перепады с помощью различных способов контрастирования. Один из наиболее простых способов заключается в вычислении дискретных разностей, что аналогично непрерывному пространственному дифференцированию. Подчеркивание вертикальных перепадов осуществляется горизонтальным (построчным)

дискретным дифференцированием. В результате формируется некоторое изображение — массив элементов

$$G(j, k) = F(j, k) - F(j, k+1). \qquad (17.4.2а)$$

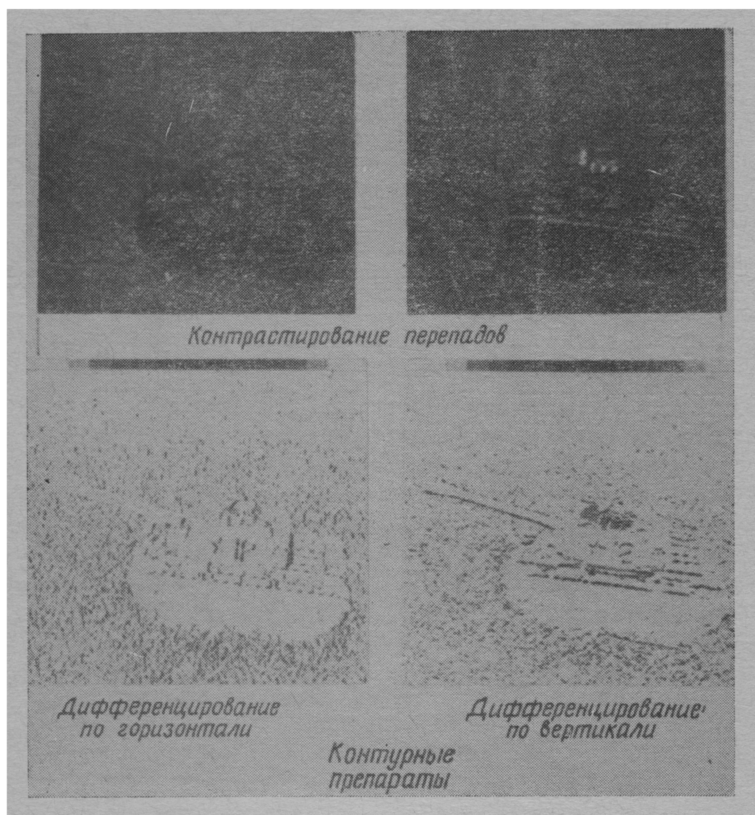


Рис. 17.4.3. Примеры выделения перепадов посредством дифференцирования по горизонтальному и вертикальному направлениям.

Аналогично осуществляется подчеркивание горизонтальных перепадов. В результате получается массив элементов

$$G(j, k) = F(j, k) - F(j+1, k). \qquad (17.4.2б)$$

Диагональное подчеркивание можно получить путем вычисления разностей уровней диагональных пар элементов изображения. На рис. 17.4.3 приведены примеры выделения перепадов посредством дискретного дифференцирования по горизонтальному и вертикальному направлениям. Контурные препараты были получены путем

сравнения вычисленных значений с порогом, величина которого соответствует 85%-ному уровню на гистограмме распределения значений модуля градиента.

Горизонтальное подчеркивание перепада также можно выполнить, вычисляя разности яркостей элементов вдоль строки изображения по формуле

$$G(j, k) = [F(j, k) - F(j, k-1)] - [F(j, k+1) - F(j, k)], \quad (17.4.3а)$$

или

$$G(j, k) = 2F(j, k) - F(j, k-1) - F(j, k+1). \quad (17.4.3б)$$

Подобные выражения существуют также для изменений яркости по вертикали и диагонали.

Двумерное дискретное дифференцирование можно выполнить с помощью операции свертки массива исходного изображения с курсовыми градиентными масками, перечисленными ниже [9, стр. 111]:

Север

$$\mathbf{H} = \begin{matrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{matrix} \quad (17.4.4а)$$

Северо-восток

$$\mathbf{H} = \begin{matrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & -1 & 1 \end{matrix} \quad (17.4.4б)$$

Восток

$$\mathbf{H} = \begin{matrix} -1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & 1 & 1 \end{matrix} \quad (17.4.4в)$$

Юго-восток

$$\mathbf{H} = \begin{matrix} -1 & -1 & 1 \\ -1 & -2 & 1 \\ 1 & 1 & 1 \end{matrix} \quad (17.4.4г)$$

Юг

$$\mathbf{H} = \begin{matrix} -1 & -1 & -1 \\ 1 & -2 & 1 \\ 1 & 1 & 1 \end{matrix} \quad (17.4.4д)$$

Юго-запад

$$\mathbf{H} = \begin{matrix} 1 & -1 & -1 \\ 1 & -2 & -1 \\ 1 & 1 & 1 \end{matrix} \qquad (17.4.4е)$$

Запад

$$\mathbf{H} = \begin{matrix} 1 & 1 & -1 \\ 1 & -2 & -1 \\ 1 & 1 & -1 \end{matrix} \qquad (17.4.4ж)$$

Северо-запад

$$\mathbf{H} = \begin{matrix} 1 & 1 & 1 \\ 1 & -2 & -1 \\ 1 & -1 & -1 \end{matrix} \qquad (17.4.4з)$$

Название курса указывает на направление склона перепада яркости, при котором маска дает максимальный отклик. Например, восточная градиентная маска дает максимальный отклик при изменении яркости в горизонтальном направлении слева направо. Следует отметить, что эти градиентные маски обладают нулевым суммарным весом, поэтому в областях изображения с постоянной яркостью они дают нулевой отклик. На рис. 17.4.4 приведены фотографии, на которых представлены изображения, контрастированные с помощью вычисления градиента, и соответствующие контурные препараты.

Повышение контраста перепадов без учета их ориентации можно получить путем свертки массива изображения с оператором Лапласа, представленным в виде маски. Несколько видов масок оператора Лапласа перечислено ниже:

Маска 1

$$\mathbf{H} = \begin{matrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{matrix} \qquad (17.4.5а)$$

Маска 2

$$\mathbf{H} = \begin{matrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{matrix} \qquad (17.4.5б)$$

Маска 3

$$\mathbf{H} = \begin{matrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{matrix}. \qquad (17.4.5в)$$

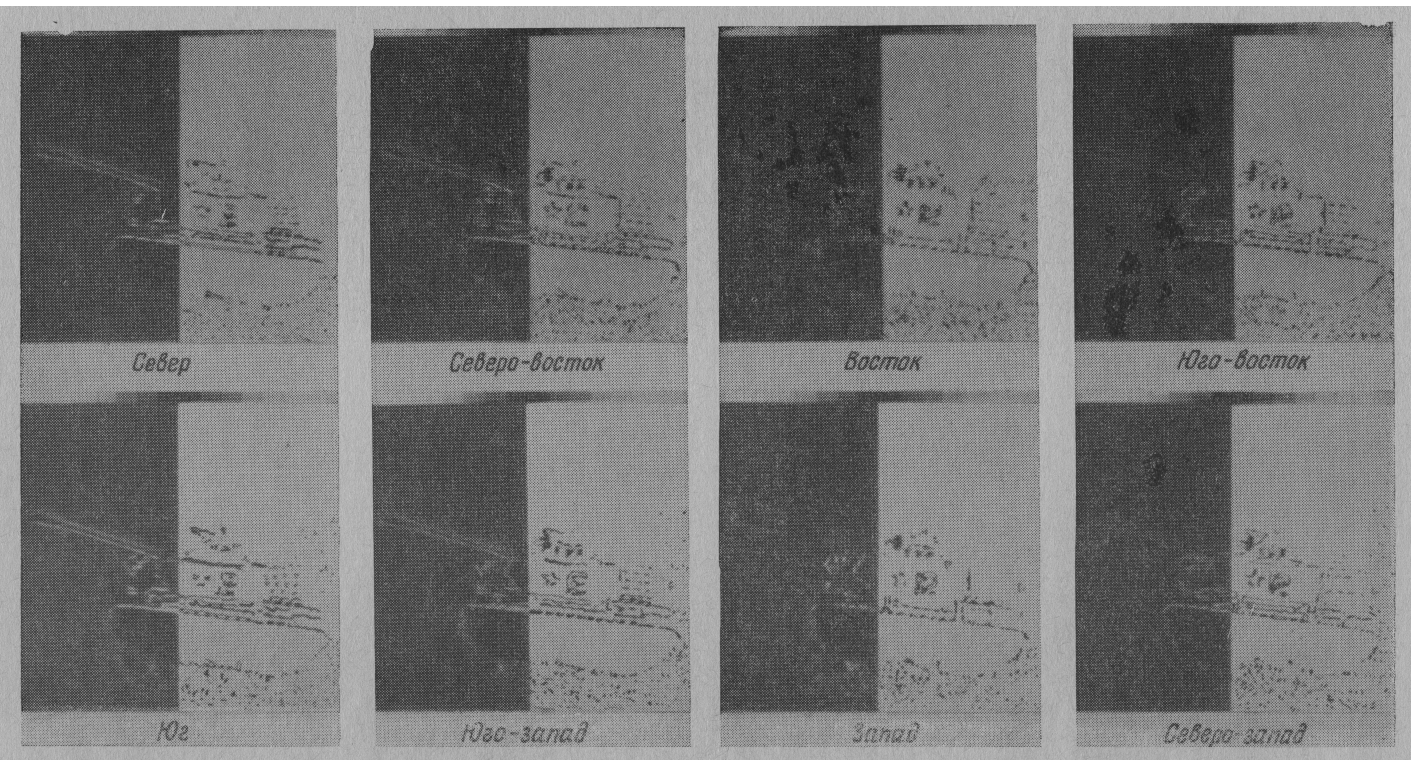


Рис. 17.4.4. Примеры выделения перепадов с помощью градиентного детектора.

Рис. 17.4.5 иллюстрирует работу детектора перепадов, выполняющего свертку массива изображения с оператором.

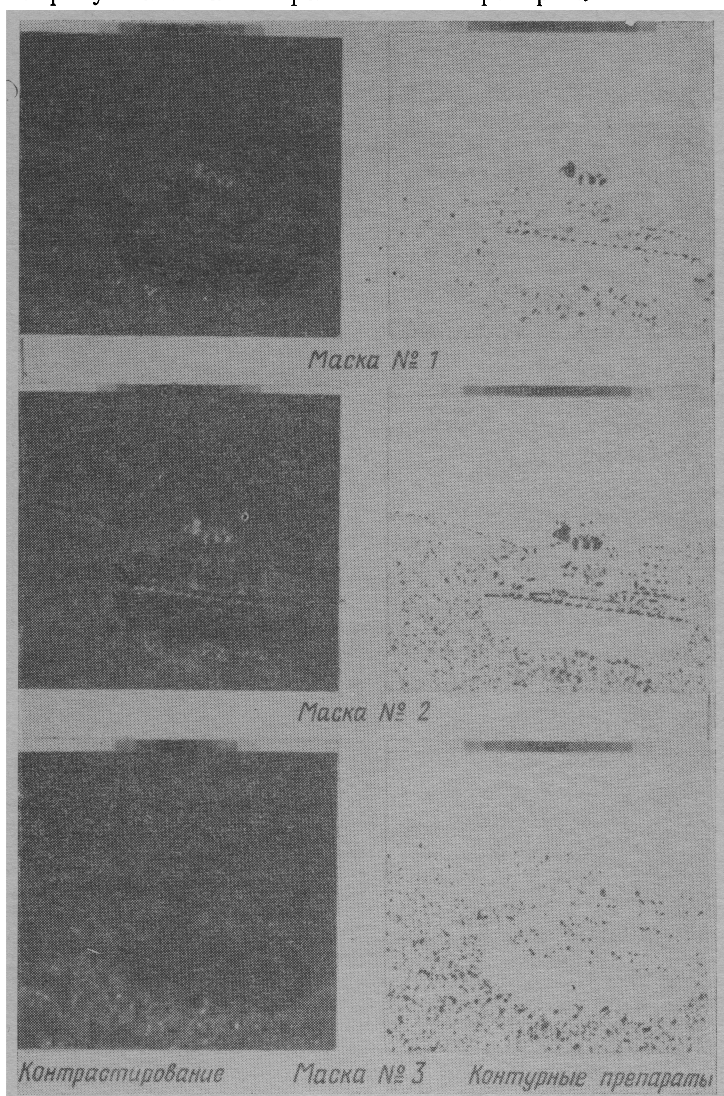


Рис. 17.4.5. Примеры выделения перепадов посредством вычисления лапласиана.

Для контрастирования перепадов можно использовать корреляционную маску, элементы которой пропорциональны коэффициентам корреляции элементов изображения [9, стр. 125]. Для

изображения, являющегося реализацией марковского процесса, корреляционная маска

$$\mathbf{H} = \begin{bmatrix} \rho_C \rho_R & -\rho_C (1 + \rho_R^2) & \rho_C \rho_R \\ -\rho_R (1 + \rho_C^2) & (1 + \rho_C^2)(1 + \rho_R^2) & -\rho_R (1 + \rho_C^2) \\ \rho_C \rho_R & -\rho_C (1 + \rho_R^2) & \rho_C \rho_R \end{bmatrix}, \tag{17.4.6}$$

где ρ_R и ρ_C — коэффициенты корреляции между соседними по строке и столбцу элементами изображения. Если $\rho_R = \rho_C = 0$, то

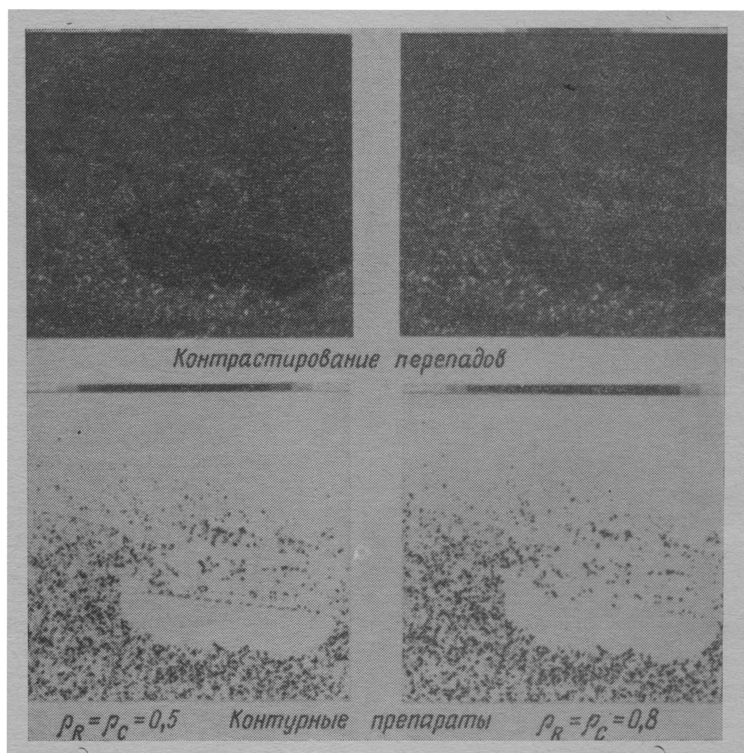


Рис. 17.4.6. Примеры выделения перепадов с помощью корреляционной маски.

корреляция между соседними элементами отсутствует и корреляционная маска не оказывает влияния на изображение; в другом частном случае, когда $\rho_R = \rho_C = 1$, корреляционная маска сводится к маске оператора Лапласа, соответствующей выражению (17.4.5в). На рис. 17.4.6 приведены примеры обнаружения перепадов с помощью корреляционной маски.

Арджаил [10] и Маклеод [11, 12] для контрастирования перепадов предложили использовать весовые функции гауссовой формы. Функция, предложенная Арджаилом, представляет собой

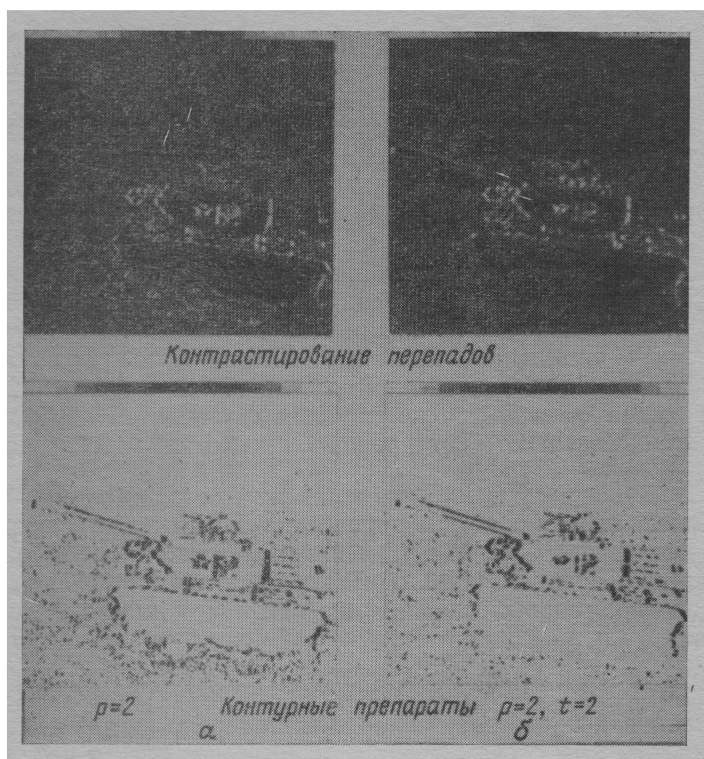


Рис. 17.4.7. Примеры выделения перепадов с помощью масок Арджаила размером 1×7 (*а*) и Маклеода размером 5×9 (*б*).

расчлененную функцию Гаусса, определяемую в одномерном случае как

$$h(x) = \exp\left\{-\frac{1}{2}\left(\frac{x}{p}\right)^2\right\}, \quad x \geqslant 0,$$

$$h(x) = -\exp\left\{-\frac{1}{2}\left(\frac{x}{p}\right)^2\right\}, \quad x < 0, \qquad (17.4.7)$$

где p — постоянная рассеяния. Функция Маклеода, которая задается выражением

$$H(x, y) = \exp\left\{-\frac{1}{2}\left(\frac{y}{t}\right)^2\right\}\left[\exp\left\{-\frac{1}{2}\left(\frac{x-p}{p}\right)^2\right\} - \exp\left\{-\frac{1}{2}\left(\frac{x+p}{p}\right)^2\right\}\right], \quad (17.4.8)$$

где p и t — постоянные рассеяния, подавляет влияние элементов изображения в области перепада яркостей и влияние других перепадов, расположенных выше и ниже основного перепада, который необходимо обнаружить. Примеры обнаружения перепадов с помощью этих масок приведены на рис. 17.4.7.

Усиление высокочастотного пространственного шума и побочные явления, связанные с операциями дискретного дифференцирования, накладывают общее ограничение на описанные выше линейные методы повышения контраста. Сглаживание шума можно ввести в линейные методы контрастирования путем линейного маскирования областей, а не отдельных элементов изображения [13]. Это достигается формированием линейной маски вида

$$H(j, k) = H_S(j, k) * H_E(i, k) \quad (17.4.9)$$

с помощью свертки одной из определенных ранее масок подчеркивания перепадов $H_E(j, k)$ с усредняющей маской $H_S(j, k)$, определяемой уравнением (12.3.2). Такое пространственное усреднение, конечно, ведет к сглаживанию как шумов, так и перепадов.

17.4.2. НЕЛИНЕЙНЫЕ МЕТОДЫ КОНТРАСТИРОВАНИЯ

В нелинейных системах обнаружения перепадов для контрастирования перед пороговым ограничением используются нелинейные комбинации значений яркости элементов изображения. В большинстве методов ограничиваются обработкой окном размером 2×2 или 3×3.

Робертс [14] для контрастирования и выделения перепадов предложил следующую простую нелинейную операцию двумерного дискретного дифференцирования:

$$G_R(j, k) = ([F(j, k) - F(j+1, k+1)]^2 + [F(j, k+1) - F(j+1, k]^2)^{1/2}. \quad (17.4.10)$$

Другая операция пространственного дифференцирования, требующая меньшего объема вычислений, определяется формулой

$$G_A(i, k) = |F(j, k) - F(j+1, k+1)| + |F(j, k+1) - F(j+1, k)|. \quad (17.4.11)$$

Можно легко показать, что

$$G_R(j, k) \leqslant G_A(j, k) \leqslant \sqrt{2}\, G_R(j, k). \qquad (17.4.12)$$

Отмечая тот из четырех элементов изображения, расположенных около обнаруженной точки, который имеет наибольшее значение

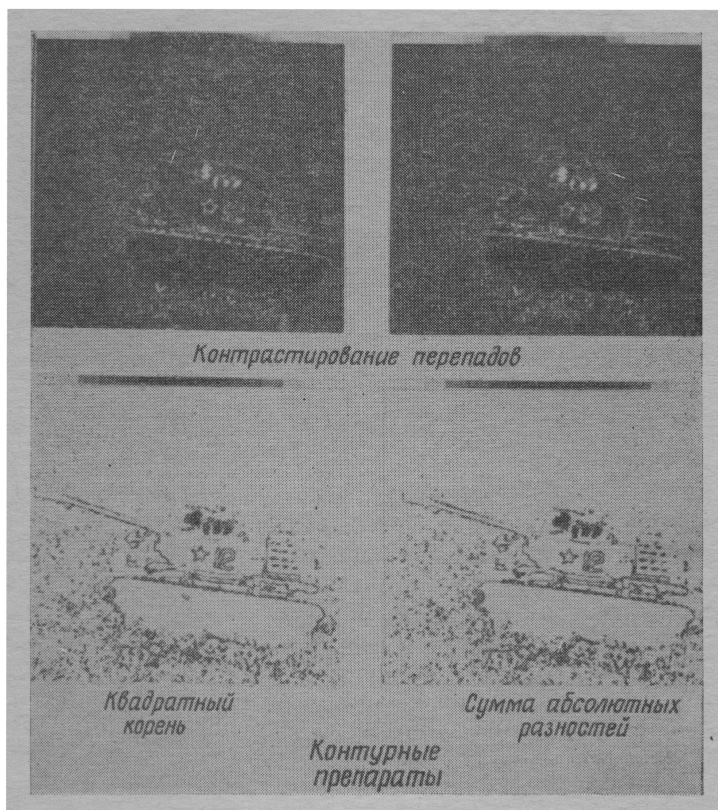


Рис. 17.4.8. Примеры выделения перепадов с помощью детекторов Робертса.

яркости, можно получить информацию о приблизительной ориентации перепада. Рис. 17.4.8 иллюстрирует работу операторов Робертса.

Собел [15, стр. 271] предложил использовать для контрастирования нелинейный оператор с окном из 3×3 элементов. На рис. 17.4.9 показаны использованные им обозначения элементов окна. Для контрастирования перепадов вычисляется величина

$$G(j, k) = \sqrt{X^2 + Y^2}, \qquad (17.4.13a)$$

где

$$X = (A_2 + 2A_3 + A_4) - (A_0 + 2A_7 + A_6), \qquad (17.4.13б)$$

$$Y = (A_0 + 2A_1 + A_2) - (A_6 + 2A_5 + A_4). \qquad (17.4.13в)$$

Величины A_i — яркости элементов окна, показанного на рис. 17.4.9. Результаты обнаружения перепадов с помощью оператора Собела приведены на рис. 17.4.10.

Другой нелинейный оператор с таким же окном из 3×3 элементов для контрастирования перепадов предложен Киршем [16]. Этот оператор определяется следующим выражением:

$$G(j,\ k) = \max\left\{1,\ \max_{i=0\div 7}\left[\left|5S_i - 3T_i\right|\right]\right\}, \qquad (17.4.14а)$$

где

$$S_i = A_i + A_{i+1} + A_{i+2}, \qquad (17.4.14б)$$

$$T_i = A_{i+3} + A_{i+4} + A_{i+5} + A_{i+6} + A_{i+7}. \qquad (17.4.14в)$$

A_0	A_1	A_2
A_7	$F(j,k)$	A_3
A_6	A_5	A_4

Рис. 17.4.9. Обозначение элементов для операторов выделения перепадов с окном размером 3×3.

Индексы у слагаемых A_i подсчитываются по модулю 8. По существу оператор Кирша дает максимальное значение курсового градиента в некоторой точке изображения без учета значения $F(j,\ k)$. Примеры обнаружения перепадов с помощью оператора Кирша приведены на рис. 17.4.11.

Уоллис [17] предложил нелинейный метод обнаружения перепадов, основанный на гомоморфной обработке изображения. Согласно этому методу, точка находится на перепаде, если величина логарифма от яркости в этой точке превосходит среднее значение логарифмов яркостей четырех ближайших соседних элементов на некоторое фиксированное значение. Элемент контрастированного изображения определяется как

$$G(j,\ k) = \log\left|F(j,\ k)\right| - \frac{1}{4}\log(A_1) - \frac{1}{4}\log(A_3) - \frac{1}{4}\log(A_5) - \frac{1}{4}\log(A_7), \qquad (17.4.15а)$$

или, что эквивалентно,

$$G(j,\ k) = \frac{1}{4}\log\left\{\frac{[F(j,\ k)]^4}{A_1A_3A_5A_7}\right\}. \qquad (17.4.15б)$$

Сравнение $G(j,\ k)$ с верхним и нижним пороговыми значениями эквивалентно сравнению дроби в скобках выражения (17.4.15б) с видоизмененным порогом. Поэтому не требуется точно вычислять

значения логарифмов. Основное преимущество логарифмического детектора перепадов кроме простоты вычислений состоит в том, что он не чувствителен к мультипликативным изменениям уровня

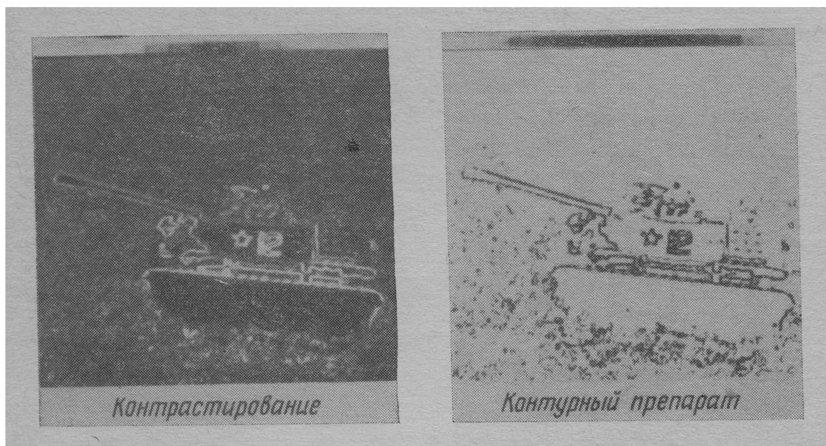


Рис. 17.4.10. Примеры выделения перепадов с помощью оператора Собела.

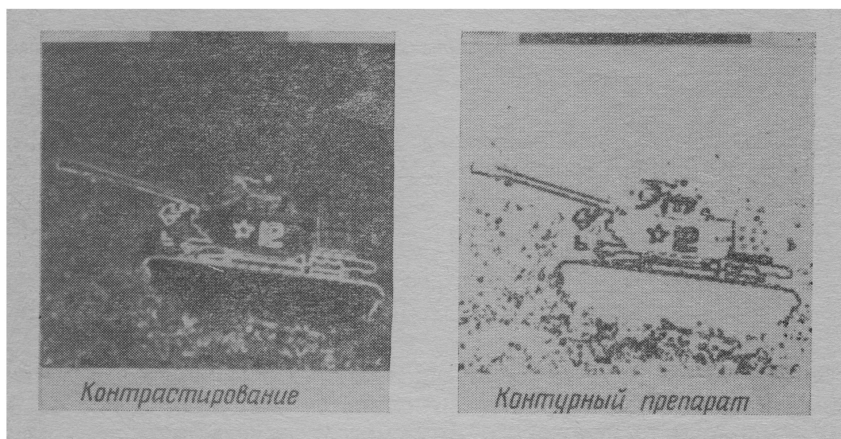


Рис. 17.4.11. Примеры выделения перепадов с помощью оператора Кирша.

яркости. Рис. 17.4.12 иллюстрирует работу логарифмического детектора перепадов.

Логарифмический метод контрастирования, определяемый выражением (17.4.15), можно рассматривать как линейное контрастирование с помощью оператора Лапласа (17.4.5а) изображения, уровни элементов которого равны логарифмам яркостей. Другие методы контрастирования также можно легко представить

в виде последовательности нелинейных поэлементных операций с дальнейшим линейным контрастированием перепадов и пороговым ограничением.

Нелинейный метод повышения контраста и выделения перепадов, основанный на вычислении произведения набора некоторых

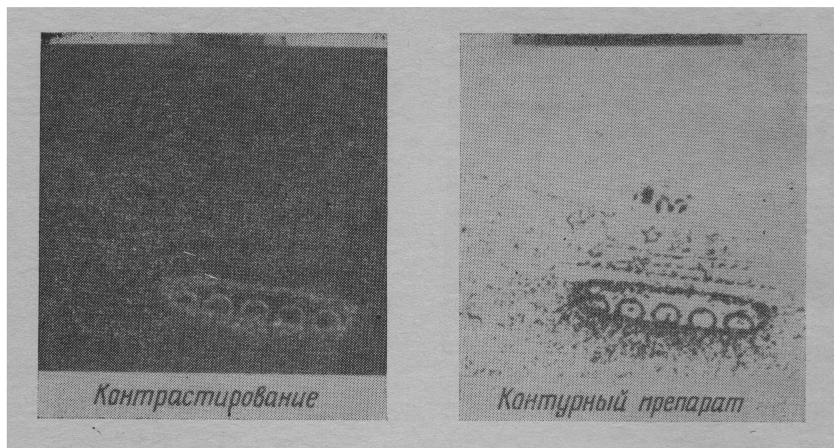


Рис. 17.4.12. Примеры выделения перепадов с помощью логарифмического оператора Лапласа.

величин, разработал Розенфельд [18]. В этом методе используется оператор построчного одномерного усреднения. Текущее среднее

$$D_M(j,\ k) = (1/M)[F(j+M-1,\ k) + F(j+M-2,\ k) + \cdots + \\ + F(j,\ k) - F(j-1,\ k) - F(j-2,\ k) - \cdots - F(j-M,\ k), \tag{17.4.16}$$

где $M = 2^m$ и m — целое, вычисляется для каждого элемента изображения. Эта операция выполняется для $M = 1, 2, 4, 8, 16, \ldots$ вплоть до некоторого желаемого верхнего предела. Затем для каждого элемента изображения вычисляется произведение

$$P_M(j,\ k) = D_1(j,\ k)\, D_2(j,\ k) \cdots D_M(j,\ k). \tag{17.4.17}$$

Понятно, что сомножители, соответствующие усреднению более высокого порядка, указывают довольно широкие пределы локализации перепада и производят некоторое подавление шумов, тогда как сомножители усреднения более низкого порядка локализуют перепад точнее, но при этом маски значительно чувствительнее к шуму. Вместе с тем можно предположить, что произведение

средних разного порядка должно обеспечивать надежное обнаружение и локализацию истинных перепадов. Обоснование этого предположения, приводимое Розенфельдом, заключается в следующем.

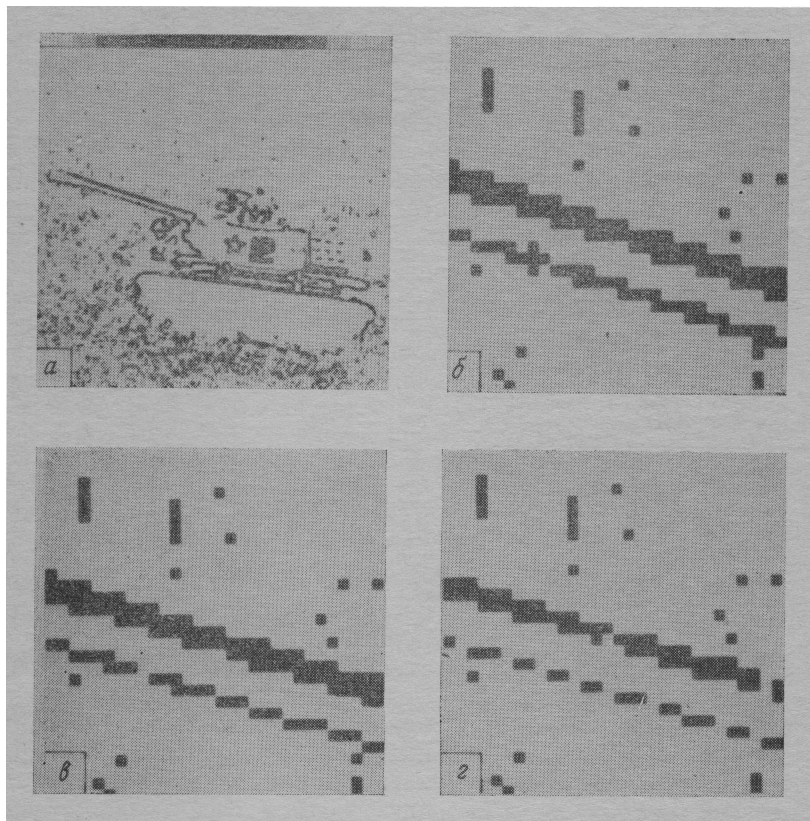


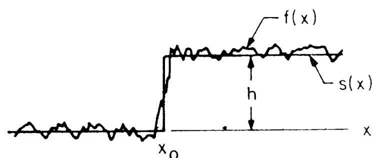
Рис. 17.4.13. Примеры обычного порогового ограничения и ограничения с подавлением доминирующими соседями для случая оператора Собела.

а — обычное пороговое ограничение; *б* — увеличенный участок изображения *а*, содержащий ствол пушки; *в* — этот же участок при коэффициенте подавления, равном 1,5; *г* — этот же участок при коэффициенте подавления, равном 1,2.

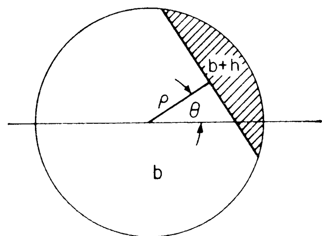
Величина $P_M(j, k)$ проявляет тенденцию обнаруживать и точно локализовать основные перепады и при этом подавлять шум. Это можно объяснить тем, что произведение велико лишь тогда, когда велики все сомножители, а по мере удаления от точки перепада сначала уменьшаются сомножители с малыми индексами, а затем и с большими.

Кроме того, Розенфельд [19] предложил нелинейную процедуру порогового ограничения для выделения больших отчетли-

вых перепадов, находящихся в окружении более мелких. Эта процедура, которую в дальнейшем будем называть «подавлением доминирующими соседями», выполняется сканированием поля контрастированных перепадов $G(j, k)$ окном малых размеров. Значение $G(j, k)$ в центре окна подавляется (приравнивается нулю) всякий раз, когда оно не оказывается наибольшим из всех отсчетов в этом окне. Затем выполняется обычное пороговое ограничение. Другой вариант такой обработки состоит в том, что подавление $G(j, k)$ разрешается лишь в случае, если в окне имеются элементы, значения которых намного превышают значения $G(j, k)$. Алгоритм подавления доминирующими соседями с последующим пороговым ограничением оказывается довольно эффективным при обнаружении перепадов, если его объединить с методом контрастирования, обеспечивающим некоторое сглаживание шума. На рис. 17.4.13 приведены примеры, иллюстрирующие работу алгоритма подавления доминирующими соседями с последующим пороговым ограничением.



а



б

Рис. 17.4.14. Аппроксимация одномерного (*а*) и двумерного (*б*) перепадов.

17.4.3. МЕТОДЫ АППРОКСИМАЦИИ ПЕРЕПАДОВ ЯРКОСТИ

Идеальные перепады можно рассматривать как одно- или двумерные сигналы, имеющие форму ступеньки (рис. 17.4.1). Тогда фрагмент реального изображения можно аппроксимировать идеальным перепадом, меняя его параметры. Если такую аппроксимацию удается сделать достаточно точно в данном месте изображения, то считается, что в этом месте имеется перепад с найденными параметрами. В случае одномерного перепада, представленного на рис. 17.4.14, *а*, функция наблюдаемого изображения $f(x)$ аппроксимируется ступенчатой функцией

$$s(x) = \begin{cases} b, & x < x_0, \\ b + h, & x \geqslant x_0. \end{cases} \tag{17.4.18}$$

Считается, что перепад существует, если среднеквадратическая ошибка аппроксимации

$$\mathscr{E} = \int_{x_0 - L}^{x_0 + L} [f(x) - s(x)]^2 \, dx \tag{17.4.19}$$

ниже некоторого порогового значения. Двумерный идеальный ступенчатый перепад определяется как

$$S(x, y) = \begin{cases} b, & x\cos\theta + y\sin\theta < \rho, \\ b+h, & x\cos\theta + y\sin\theta \geqslant \rho, \end{cases} \tag{17.4.20}$$

где θ и ρ — полярные координаты точки перепада, ближайшей к центру исследуемой круговой области. Ошибка аппроксимации вычисляется по формуле

$$\mathscr{E} = \iint [F(x, y) - S(x, y)]^2\, dx\, dy. \tag{17.4.21}$$

Хюккель [20] разработал процедуру аппроксимации двумерного перепада, при которой фрагмент изображения, оказавшийся внутри круга (рис. 17.4.14, *б*), раскладывается по набору двумерных базисных функций в ряд Фурье в полярных координатах. Пусть $H_i(x, y)$ — базисные функции. Тогда коэффициенты разложения для изображения и идеального ступенчатого перепада будут иметь вид



Рис. 17.4.15. Пример выделения перепадов с помощью оператора Хюккеля [21].

$$a_i = \iint H_i(x, y)\, F(x, y)\, dx\, dy, \tag{17.4.22а}$$

$$b_i = \iint H_i(x, y)\, S(x, y)\, dx\, dy. \tag{17.4.22б}$$

Следует отметить, что $S(x, y)$ определяется в параметрической форме в виде набора параметров (b, h, ρ, θ). В алгоритме Хюккеля разложение ограничено восемью базисными функциями с целью сокращения объема вычислений и обеспечения некоторого сглаживания шума. Минимизация среднеквадратического отклонения (17.4.21) эквивалентна минимизации величины $(a_i - b_i)^2$ для всех коэффициентов. Хюккель осуществил минимизацию, прибегнув к некоторым упрощениям, и вывел систему нелинейных уравнений, посредством которых параметры перепада $(\hat{b}, \hat{h}, \hat{\rho}, \hat{\theta})$ выражаются через коэффициенты a_i. После минимизации проводится сравнение функции $F(x, y)$, описывающей исходное изображение, с ее аппроксимацией. Если результаты сравнения оказываются неудовлетворительными, то считается, что в данной окрестности

перепада нет. Если же аппроксимация достаточно точна, то значение найденного контраста h сравнивается с заданным порогом.

Сложность алгоритма Хюккеля затрудняет его теоретический анализ. Однако данные экспериментов показывают, что оператор Хюккеля работает достаточно хорошо в качестве детектора перепадов даже на зашумленных изображениях и изображениях с сильно выраженной текстурой [21]. Рис. 17.4.15 иллюстрирует обнаружение перепадов при помощи алгоритма Хюккеля.

17.4.4. СТАТИСТИЧЕСКИЕ МЕТОДЫ ОБНАРУЖЕНИЯ ПЕРЕПАДОВ

Основной способ контрастирования перепадов с последующим пороговым ограничением, разнообразные формы которого были рассмотрены выше, по своему характеру является эвристическим. Каждый алгоритм необходимо «настраивать» методом проб и ошибок с помощью набора испытательных картинок. Обычно получаемый при этом результат можно оценить лишь экспериментально, но нельзя предсказать заранее. Метод аппроксимации, хотя и сформулирован математически, имеет множество таких же недостатков. Кроме того, эти методы не справляются достаточно эффективно с шумом. Рассмотрим теперь класс методов обнаружения перепадов, основанных на теории статистического обнаружения [1]).

Методы контрастирования и аппроксимации для задач обнаружения перепадов при наличии шума можно улучшить путем введения способов обработки, основанных на статистических свойствах шума. Рассмотрим общий случай, когда наблюдаемое изображение представляет собой массив элементов $Y(j, k) = F(j, k) + N(j, k)$, где $F(j, k)$ и $N(j, k)$ — элементы исходного изображения и шумовой составляющей. При отсутствии шума линейные методы контрастирования дают массив элементов контрастированного изображения $G(j, k) = H(j, k) * F(j, k)$, где $H(j, k)$ — импульсный отклик контрастирующей системы. При наличии шума контрастированное изображение, представляемое массивом $G_Y(j, k) = Y(j, k) * H(j, k)$, содержит шумовую составляющую $H(j, k) * N(j, k)$, которая увеличивается с возрастанием пространственных частот. Эту составляющую можно значительно уменьшить, формируя винеровскую оценку $\hat{G}(j, k)$ элемента контрастированного изображения по наблюдаемой реализации — массиву элементов $Y(j, k)$, как было описано в части 4. Однако поскольку процесс получения винеровской оценки линеен, то $\hat{G}(j, k) = H(j, k) * \hat{F}(j, k)$, где $\hat{F}(j, k)$ — оценка идеального изображения $F(j, k)$ по массиву $Y(j, k)$. Таким образом, можно

---

[1]) Подход с использованием теории статистического обнаружения ранее применялся при обнаружении «деталей» в телевизионном изображении [22, 23].

найти оценку $\hat{F}(j, k)$, а затем контрастировать перепады. При этом следует соблюдать осторожность, поскольку оптимальное оценивание до контрастирования приводит к тому же результату, что и оптимальное оценивание контрастированного изображения, только в случае линейных процессов. Если пороговому ограничению подвергается модуль $G_Y(j, k)$, то среднеквадратическая оценка величины $F(j, k)$ до контрастирования не обязательно даст оптимальный результат. Однако, по-видимому, применение оценки до контрастирования приведет к лучшим результатам, чем контрастирование без предварительной оценки.

Если идеальный перепад можно описать некоторой детерминированной моделью, например, такой, как показана на рис. 17.4.1, то для обнаружения перепадов можно применить методы согласованной фильтрации. Этот вопрос рассматривается в гл. 19.

17.5. ЭФФЕКТИВНОСТЬ АЛГОРИТМОВ ОБНАРУЖЕНИЯ ПЕРЕПАДОВ ЯРКОСТИ

В литературе описано сравнительно мало исследований по оценке эффективности детекторов перепада [24, 25]. Эффективность трудно оценить вследствие большого числа предложенных методов, из-за трудностей в определении наилучших параметров, связанных с каждым методом, а также из-за отсутствия единого критерия эффективности.

При разработке критерия эффективности для детекторов перепада целесообразно делать различие между обязательной и вспомогательной информацией, которую необходимо получить с помощью таких детекторов. Очевидно, совершенно необходимо определять положение перепада. Другая представляющая интерес информация включает высоту и крутизну перепада яркости, а также его пространственную ориентацию. Полезной информацией является также достоверность решения о перепаде, например о степени соответствия фрагмента истинного изображения модели идеального перепада. К сожалению, имеется немного детекторов, дающих этот полный набор данных.

Существуют три основных типа ошибок, связанных с определением положения перепадов яркости: 1) пропуск истинных перепадов; 2) ошибка в определении их положения; 3) принятие шумовых выбросов за перепад (ложное обнаружение). На рис. 17.5.1 показаны типичный фрагмент дискретного изображения, содержащий перепад яркости, идеальный перепад и контурные препараты при различных типах ошибок.

Общая стратегия в задачах обнаружения сигналов заключается в установлении некоторой границы вероятности ложного обнаружения, являющегося следствием воздействия шума, с последующей максимизацией вероятности правильного обнаружения.

Распространение этой идеи на обнаружение перепадов ведет к установлению порога обнаружения на таком уровне, при котором вероятность ложного обнаружения, обусловленного одним лишь шумом, не превосходит некоторого заданного значения. Вероятность обнаружения истинного перепада можно легко вычислить, сравнивая контурные изображения, полученные с помощью идеального и реального детекторов перепада. Определение величины штрафа за ошибки в локализации перепадов представляет собой более трудную задачу. Ясно, что детекторы, дающие «смазанный» перепад, должны быть «оштрафованы». Следует отдавать предпочтение детекторам, которые локализуют перепады, хотя и дают при этом небольшую ошибку. Точность в определении положения перепада можно оценить величиной

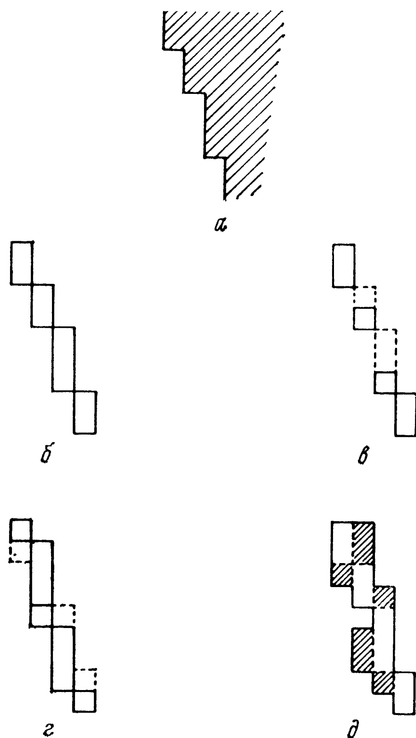


Рис. 17.5.1. Виды ошибок при определении положения перепадов.

а — фрагмент изображения; *б* — идеальное обнаружение; *в* — дробление контура; *г* — смещение контура; *д* — смазанный контур.

$$R = (1/I_N) \sum_{i=1}^{I_A} 1/(1 + \alpha d^2), \qquad (17.5.1)$$

где $I_N = \max(I_I, I_A)$, а I_I и I_A — число точек перепадов в идеальном и реальном контурных препаратах, α — масштабный множитель, d — расстояние между точкой действительного перепада и линией, состоящей из точек идеального перепада, измеренное по нормали к этой линии. Значение критерия (17.5.1) нормализовано так, что $R = 1$ для точно выделенного перепада. Масштабный множитель можно подобрать таким образом, чтобы установить штрафы для перепадов, истинное положение которых отличается от обнаруженного. Множитель $1/I_N$ обеспечивает штраф за смазанные или разбитые контуры. Например, если $\alpha = {}^1/_9$, то при обнаружении вертикального перепада, отстоящего на один элемент изображения относительно истинного положения $R = 0{,}90$, сдвиг на два элемента приводит к значению $R = 0{,}69$. При $\alpha =$

$= 1/_9$ смазанный контур шириной в три элемента изображения, центр которого совпадает с центром истинного вертикального перепада, дает значение $R = 0{,}93$, а при ширине в пять элементов

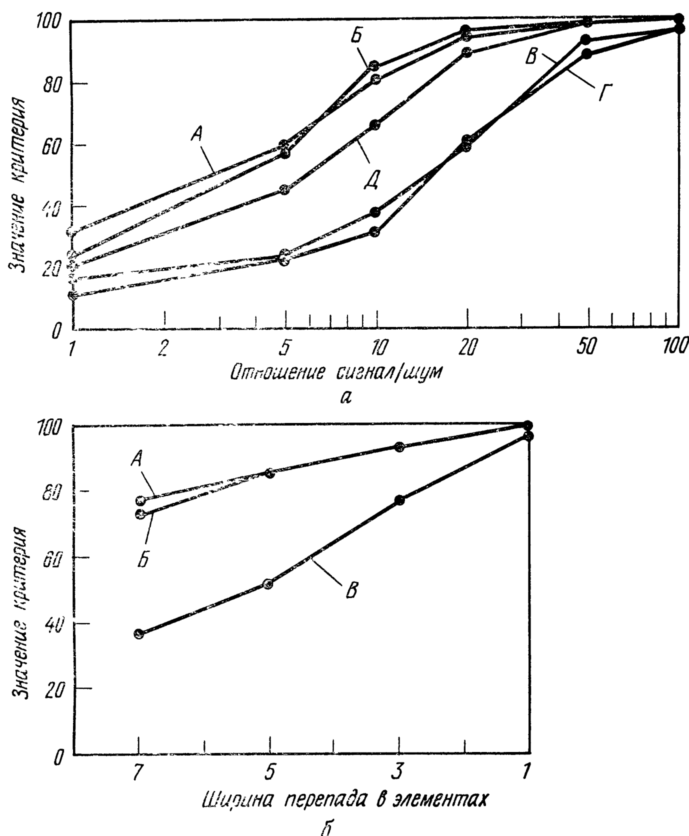


Рис. 17.5.2. Графики зависимости величины критерия эффективной локализации перепадов от отношения сигнал/шум (С/Ш) и ширины перепада для различных детекторов.

а — зависимость критерия от С/Ш ($h = 25$, $W = 1$)
А — оператор Кирша, *Б* — оператор Собела; *В* — оператор Робертса с квадратным корнем; *Г* — оператор Робертса с суммой абсолютных разностей; *Д* — восточный курсовой оператор.
б — зависимость критерия от ширины параметра ($h = 25$, С/Ш $= 100$).
А — оператор Собела, *Б* — восточный курсовой оператор; *В* — оператор Робертса с суммой абсолютных разностей

$R = 0{,}84$. Более высокое значение R для смазанного контура по сравнению со значением для сдвинутого перепада представляется разумным, поскольку имеется возможность путем последующей обработки сделать смазанный контур более тонким,

Некоторые детекторы позволяют определять высоту, крутизну и ориентацию перепада. Для этих детекторов полезно определить среднеквадратическое отклонение этих параметров от их истинных значений, усредненное по всем истинным точкам перепадов, которые совпадают с обнаруженными точками.

Методика получения оценки эффективности, описанная выше, была применена к некоторым наиболее обещающим способам обнаружения перепадов, рассмотренным в разд. 17.4. При этом использовалось испытательное изображение, состоящее из 64×64 элементов, яркость которых менялась в диапазоне 0 — 255. В центре этого изображения находился вертикально ориентированный перепад яркости с переменными контрастом и крутизной наклона. На изображение перепада был наложен независимый гауссов шум со среднеквадратическим отклонением σ_n. Результирующая картинка была ограничена по уровню в соответствии с диапазоном яркостей (0 — 255). Отношение сигнал/шум определялось как

$$\text{С/Ш} = h^2/\sigma_n^2, \tag{17.5.2}$$

где h — высота перепада. Поскольку цель исследования заключалась в сравнении эффективности различных методов обнаружения перепадов, каждый детектор обязательно должен быть рассмотрен с точки зрения его наилучших возможностей. Поэтому, прежде чем их сравнивать, разрешалось настраивать каждый детектор как на полях со случайным шумом без перепадов, так и на реальных испытательных изображениях. Для каждого детектора было установлено значение порога, при котором достигалось максимальное значение величины R при максимально допустимой частоте ложных обнаружений.

На рис. 17.5.2, *а* приведены графики зависимости величины R от отношения сигнал/шум для нескольких детекторов при $\alpha = 1/9$. Зависимость величины R от ширины контура показана на рис. 17.5.2, *б*. Из этих графиков видно, что, как и следовало ожидать, «качество» обнаружения мало для слабоконтрастных широких зашумленных перепадов и велико в противном случае. Некоторые методы обнаружения обладают лучшими показателями для всех испытательных изображений. Для субъективной проверки обоснованности критерия качества обнаружения на рис. 17.5.3 и 17.5.4 представлены контурные препараты, полученные для нескольких изображений с помощью детекторов с высоким и низким качеством обнаружения перепада. Эти фотографии подтверждают практическую пригодность данного критерия. Высокое качество обнаружения обычно соответствует хорошему визуальному определению местоположения перепада, и наоборот.

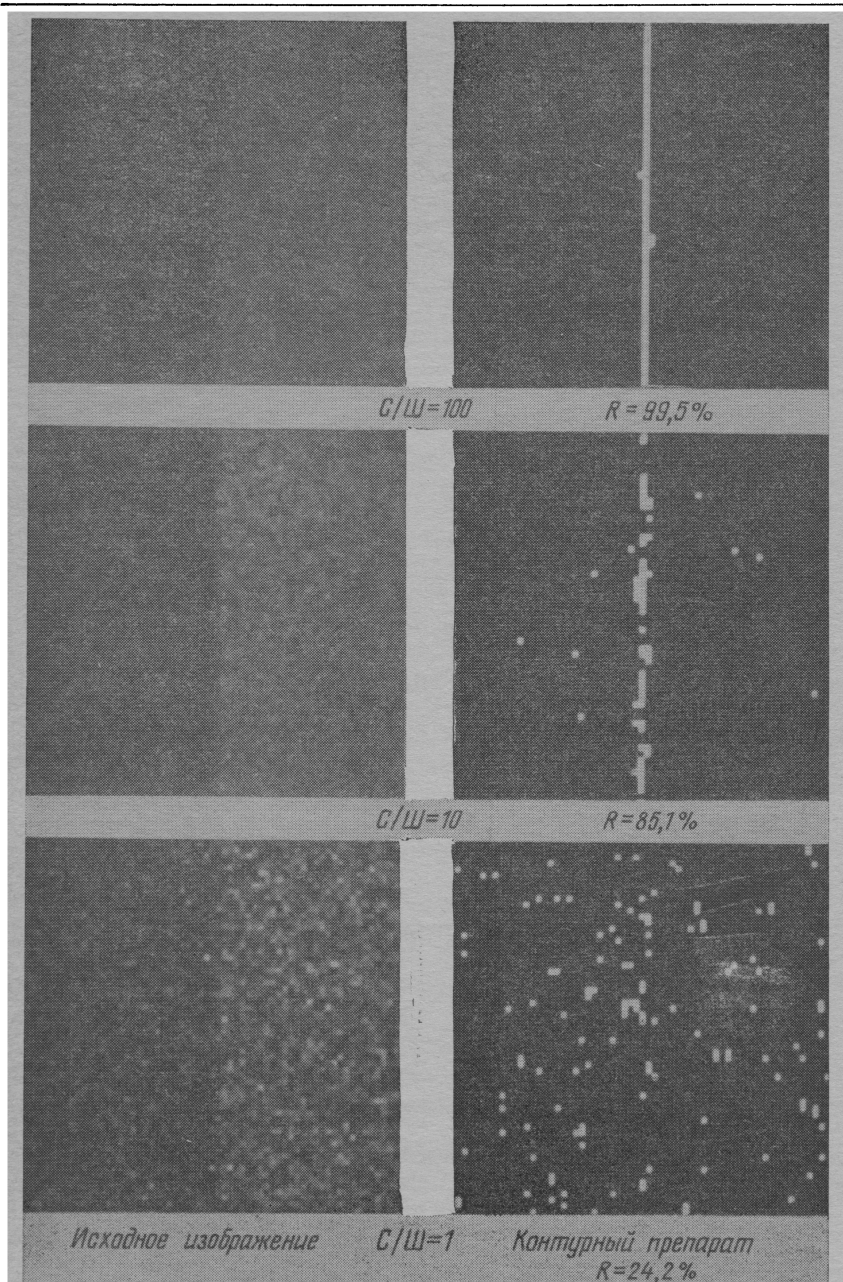


Рис. 17.5.3. Определение положения перепада с помощью оператора Собела при разных отношениях сигнал/шум, $h = 25$, $W = 1$, $\alpha = {}^1/_9$.

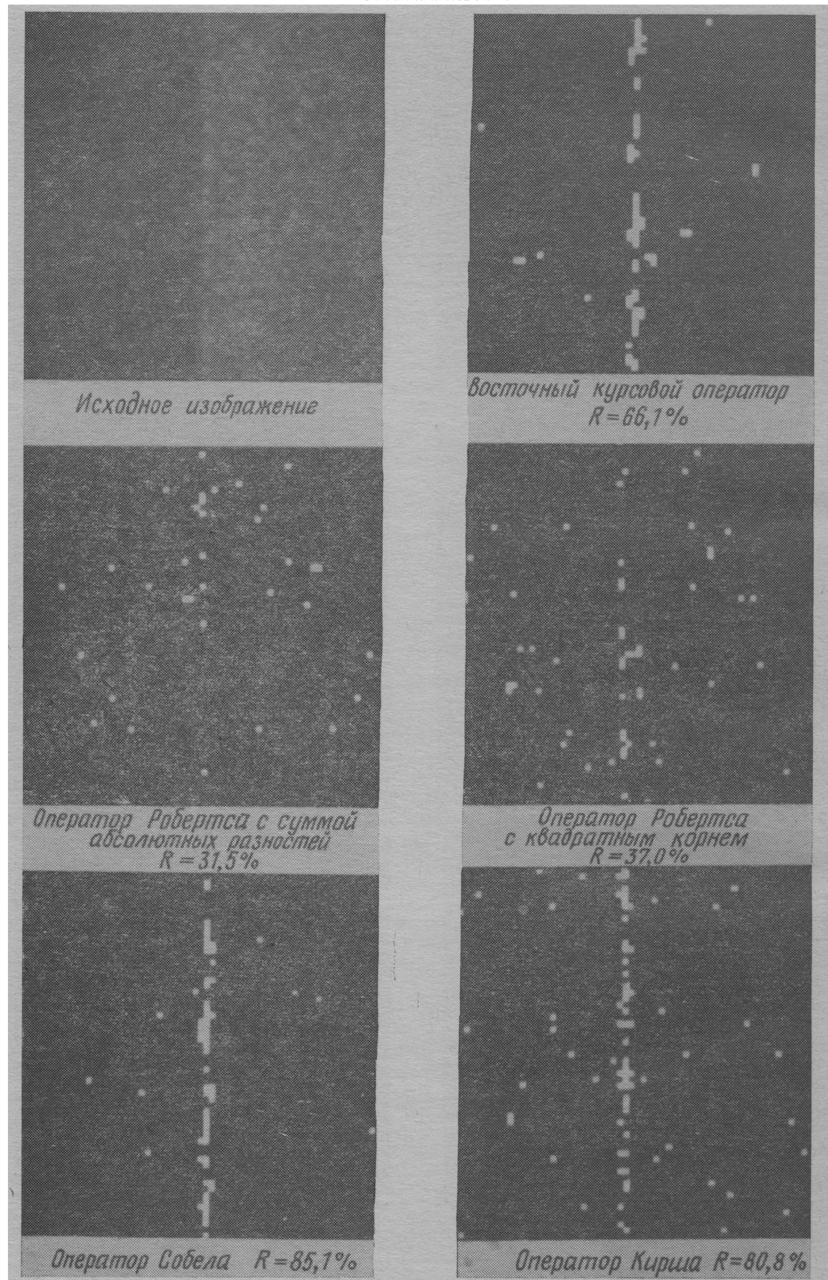


Рис. 17.5.4. Определение положения перепада различными детекторами при С/Ш = 10, $h = 25$, $W = 1$, $\alpha = {}^1/_9$.

17.6. ЦВЕТОВЫЕ КОНТУРЫ

Цветные изображения можно описывать количественно, определяя для каждого элемента координаты цвета T_1, T_2, T_3, пропорциональные количеству красного, зеленого и синего основных цветов, необходимых для уравнивания цвета этого элемента. Яркость элемента представляет собой взвешенную сумму $Y = \alpha_1 T_1 + \alpha_2 T_2 + \alpha_3 T_3$ значений координат цвета, где α_i — весовые коэффициенты.

Существует несколько определений цветовых контуров [26, 27]. Можно считать, что контур на цветном изображении соответствует только перепаду яркости. Однако это определение исключает контуры, связанные с резкими изменениями цветового тона и насыщенности в областях с постоянной яркостью. Другой подход состоит в том, что исследуется каждая из трех компонент изображений, и считается, что цветовой контур существует, если яркостный перепад обнаружен одновременно на всех компонентах. Третье определение состоит в том, что контур имеет место, если расстояние между векторами цвета по обе стороны предполагаемого перепада превосходит некоторый пороговый уровень. В случае двух последних определений цветовых контуров результаты обнаружения полностью зависят от выбранной системы координат цвета. Хорошо известно, что равные изменения значений красной, зеленой и синей координат воспринимаются не одинаково. В гл. 3 были рассмотрены системы координат *Lab* и $U^*V^*W^*$, которые представляют собой нелинейные преобразования исходных координат цвета. Эти системы координат обладают близкой к равномерной метрикой, и, следовательно, их целесообразно использовать для анализа контуров на цветных изображениях.

17.7. ПРИЗНАКИ ПЯТНА И ЛИНИИ

Пятном называется относительно маленькая область изображения, яркость которой значительно отличается от яркости ее окрестности. Отрезок линии определяется распределением яркости, которое должно иметь U-образную (или ∩-образную) форму в направлении, перпендикулярном направлению линии, и быть вытянутым вдоль линии.

Розенфельд и др. [28—30] предложили алгоритм для обнаружения пятна, при использовании которого сначала осуществляется сглаживание изображения с помощью маски размером $W \times W$ элементов, представляющей собой фильтр нижних частот. Затем уровень каждого элемента усредненного изображения сравнивается со средним значением уровней его четырех соседей (справа, слева, сверху и снизу), удаленных от центрального элемента на расстояние в W элементов. Пятно выделяется, если эта раз-

ность достаточно велика. Другой подобный подход [31] предусматривает формирование разности между средней яркостью в окне размером $W \times W$ и средней яркостью окружающей кольцевой области шириной W. Оба этих метода представляют собой особые случаи линейного контрастирования пятна с последующим пороговым ограничением, когда контрастирующая маска есть составная маска вида

$$H(j, k) = H_S(j, k) * H_L(j, k), \qquad (17.7.1)$$

где $H_S(j, k)$ — сглаживающая маска, представляющая собой фильтр нижних частот, определяемый уравнением (12.3.2), а $H_L(j, k)$ — одна из масок оператора Лапласа (17.4.5).

Можно предполагать, что линии и полосы являются локальными фрагментами более крупных структур. Говорят, что через некоторую точку изображения проходит линия, если вдоль трех прямых, перпендикулярных линии и проходящих через эту точку и две ближайшие соседние точки, лежащие на линии, яркость меняется от низкого уровня к более высокому, а затем снова к более низкому (или наоборот) [29]. Один из подходов к выделению линии состоит в том, что выполняется пространственное объединение откликов детектора перепада внутри локальной окрестности. Эта задача исследуется в гл. 18. Линии единичной ширины можно обнаружить путем свертки изображения с набором линейных масок

$$\mathbf{H}_1 = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}, \qquad (17.7.2а)$$

$$\mathbf{H}_2 = \begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}, \qquad (17.7.2б)$$

$$\mathbf{H}_3 = \begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix}, \qquad (17.7.2в)$$

$$\mathbf{H}_4 = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}, \qquad (17.7.2г)$$

с помощью которых выполняется контрастирование вертикальных, горизонтальных и диагональных отрезков линий. Пороговое ограничение контрастированных изображений приводит к обнаружению местоположения отрезка линии. Хюккель [30] обобщил идею аппроксимации перепада и предложил выполнять обнаружение отрезков линий внутри области круглой формы. Фрагмент изоб-

ражения внутри круглого окна аппроксимируется идеальным отрезком линии, который имеет вид узкой полоски, ограниченной параллельными прямыми; яркость внутри этой полоски отличается от яркости соседней области. Если удается добиться достаточно точной аппроксимации, то считается, что в изображении есть отрезок линии с подобранными в процессе аппроксимации параметрами.

17.8. ТЕКСТУРНЫЕ ПРИЗНАКИ

Бóльшая часть естественных сцен лишена существенных деталей на больших пространствах. В этих областях сцену часто можно охарактеризовать как проявление повторяющейся структуры, аналогичной структуре ткани или узору кафельного пола. Можно привести много примеров, когда желательно определять границы текстурной области и размер зерна текстуры внутри каждой области. Первая задача рассматривается в разд. 18.7. Настоящий раздел посвящен рассмотрению количественного описания текстуры.

Некоторые исследователи пытались дать качественное определение текстуры. Пикетт [31] дал такую формулировку: «текстура используется для описания двумерных массивов изменений яркости... Элементы текстуры и правила их пространственной организации или расположения можно произвольно менять, если только остаются неизменными характеристики повторяемости изменений яркости». Хоукинс [32] дал более подробное описание текстуры: «По-видимому, текстура охватывает следующие свойства изображения: 1) в нем можно найти фрагмент, «рисунок» которого регулярно повторяется в пределах области, которая велика по сравнению с размером фрагмента; 2) этот «рисунок» образуется элементарными составными частями фрагмента, размещенными в некотором неслучайном порядке; 3) элементарные части — это примерно однородные единицы, имеющие приблизительно одинаковую форму во всей текстурной области». Хотя эти описания текстуры кажутся разумными, они не приводят непосредственно к простым количественным признакам текстуры в том смысле, в котором представление о резком перепаде яркости ведет к количественному его определению через параметры, характеризующие положение перепада в пространстве, крутизну и высоту.

Текстуру можно разделить на искусственную и естественную. Искусственные текстуры — это структуры из графических знаков, расположенных на нейтральном фоне. Такими знаками могут быть отрезки линии, точки, звездочки или буквы и цифры. Несколько примеров искусственных текстур представлено на рис. 17.8.1. Естественные текстуры, как подразумевается в их названии, это изображения естественных сцен, содержащие почти периодические

структуры. Примерами могут служить фотографии кирпичных стен, черепицы крыш, песка, травы и т. д. Бродац [33] опубликовал альбом текстур, встречающихся в природе. На рис. 17.8.2 приве-

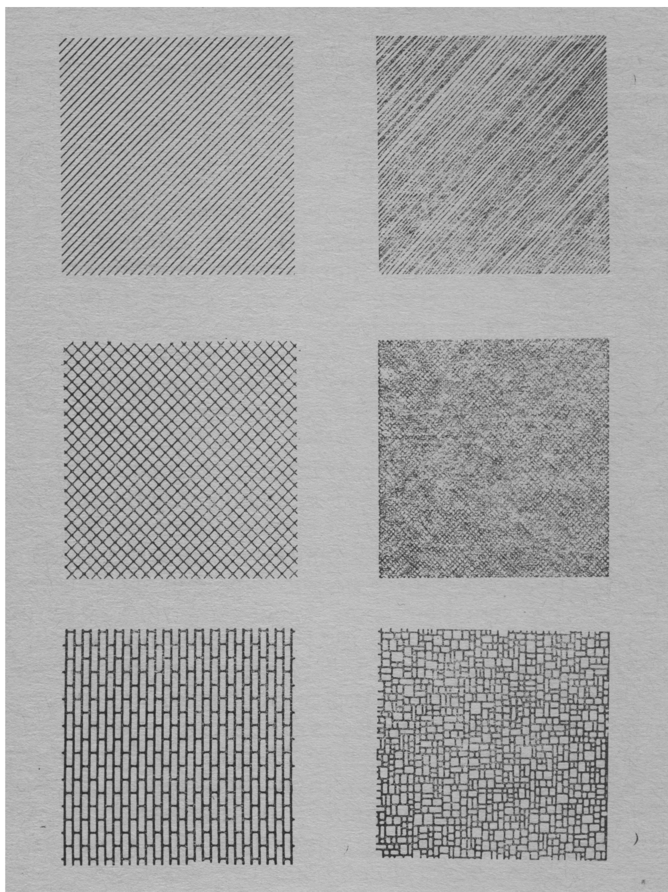


Рис. 17.8.1. Примеры искусственных текстур [33].

дено несколько примеров естественных текстур. Дальнейший анализ текстур ограничен естественными текстурами.

Текстура часто описывается качественно размером ее зерна. Например, кусок шерстяной ткани «грубее» куска шелковой ткани при одинаковых условиях наблюдения. Размер зерна связан с периодом пространственной повторяемости локальной структуры. Большой период соответствует крупной текстуре, а малый — мелкой. Ясно, что размер зерна недостаточен для количественного измерения текстуры, но его можно по крайней мере использовать

для оценки направления, в котором должны изменяться текстурные признаки, т. е. малые численные значения признаков должны соответствовать мелкой текстуре, а большие значения — крупной. Нужно учитывать, что текстура — это свойство окрестности точки

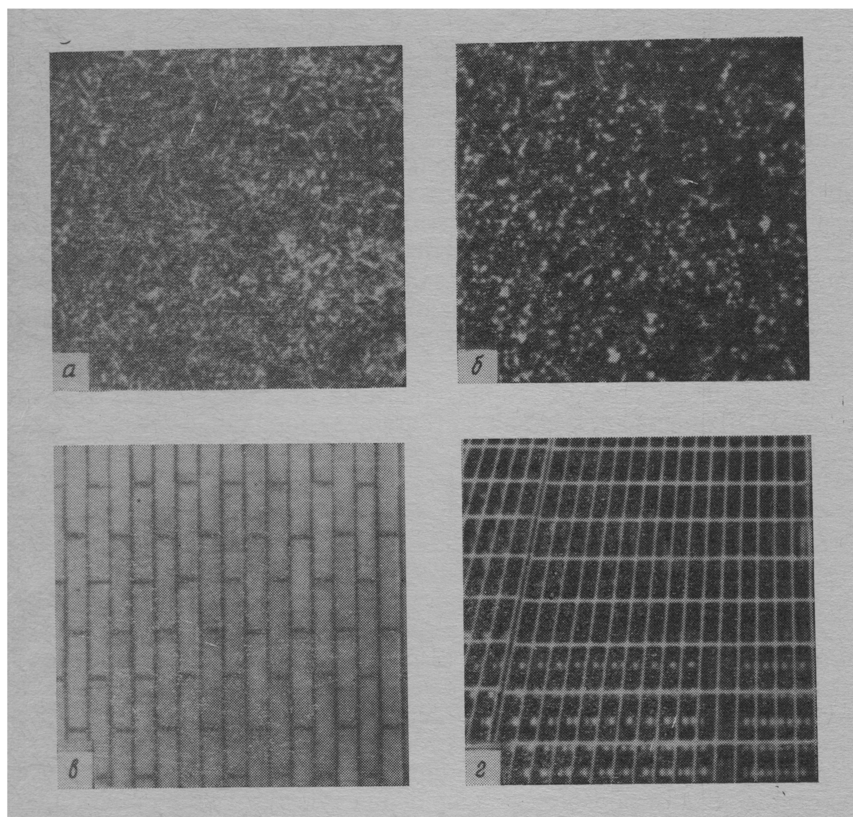


Рис. 17.8.2. Примеры естественных текстур.
а — трава, *б* — плющ, *в* — кирпичная кладка; *г* — решетка.

изображения. Следовательно, текстурные признаки по своей природе зависят от размера окрестности, на которой они определяются. Поскольку текстура — пространственное свойство, измерения ее признаков должны быть ограничены областями, обладающими относительной однородностью. Поэтому, прежде чем делать попытки измерить текстуру, необходимо установить границы области однородной текстуры путем наблюдения или с помощью одного из методов автоматической сегментации изображения, описанных в гл. 18.

Несколько исследований [4, 34, 35] посвящено анализу текстур с помощью спектра Фурье (см. разд. 17.3). Поскольку размер зерна текстуры пропорционален пространственному периоду, область крупнозернистой текстуры должна давать спектр Фурье, энергия которого сосредоточена на низких пространственных частотах. Наоборот, для областей мелкозернистой текстуры энергия спектра концентрируется на высоких пространственных частотах. Хотя такое соответствие отчасти и существует, часто возникают трудности, связанные с пространственным изменением периода и фазы повторений структуры. Эксперименты [32, 34] показали, что существует значительное перекрытие спектров для областей с заметно отличающейся естественной текстурой, например, таких, как городские кварталы, сельские районы и лесные массивы, выделенные на аэрофотоснимках. С другой стороны, спектральный фурье-анализ оказался успешным [7, 8] при обнаружении и классификации антракоза легких у горняков, который визуально проявляется в виде диффузных текстурных отклонений снимков легких от нормы.

В качестве основной характеристики текстуры была предложена [36] пространственная автокорреляционная функция. Рассмотрим эту функцию

$$A(\varepsilon, \eta; j, k) = \frac{\sum\limits_{m=j-W}^{j+W} \sum\limits_{n=k-W}^{k+W} F(m, n) F(m-\varepsilon, n-\eta)}{\sum\limits_{m=j-W}^{j+W} \sum\limits_{n=k-W}^{k+W} [F(m, n)]^2}, \qquad (17.8.1)$$

вычисляемую на окне размером $(2W + 1) \times (2W + 1)$ для каждой точки изображения (j, k) и при смещениях $\varepsilon, \eta = 0, \pm 1, \pm 2, \ldots, \pm T$. Предполагается, что при фиксированном сдвиге (ε, η) области крупнозернистой текстуры будет соответствовать более высокое значение корреляционной функции, чем области мелкозернистой текстуры. Таким образом, размер зерна текстуры пропорционален ширине автокорреляционной функции. Одной из возможных мер ширины автокорреляции является второй момент

$$T(j, k) = \sum_{\varepsilon=-T}^{T} \sum_{\eta=-T}^{T} \varepsilon^2 \eta^2 A(\varepsilon, \eta; j, k). \qquad (17.8.2)$$

Розенфельд и Трой [36] предложили в качестве текстурного признака использовать число перепадов яркости в окрестности точки. Сначала с помощью некоторой системы обнаружения перепадов создается контурный препарат $E(j, k)$, такой, что $E(j, k) = 1$ для обнаруженной точки перепада и $E(j, k) = 0$ в противном случае. Обычно порог обнаружения устанавливается ниже, чем в случае выделения граничных точек областей постоянной яркости.

Затем формируется текстурный признак

$$T(j,\ k) = [1/(2W+1)^2] \sum_{m=j-W}^{j+W} \sum_{n=k-W}^{k+W} E(m,\ n), \qquad (17.8.3)$$

вычисляемый на окне размером $(2W+1) \times (2W+1)$ для каждой рассматриваемой точки изображения.

Харалик, Шанмуган и Динштейн [2] предложили ряд текстурных признаков, основанных на свойствах гистограммы распределения частот совместных значений яркости пары элементов изображения. Если на изображении имеется область с мелкой текстурой, то эта гистограмма будет близка к однородной, а для грубой текстуры окажется сосредоточенной вдоль диагонали. Рассмотрим пару элементов изображения $F(j,\ k)$ и $F(m,\ n)$, расстояние между которыми определяется вектором с модулем r и углом θ относительно горизонтальной оси. Пусть $P(a,\ b;\ j,\ k,\ r,\ \theta)$ — распределение частот, измеренное в окне размером $(2W+1) \times (2W+1)$, где значения яркости квантованы в диапазоне $0 \leqslant a,\ b \leqslant L-1$. Это распределение можно рассматривать как оценку совместного распределения вероятностей

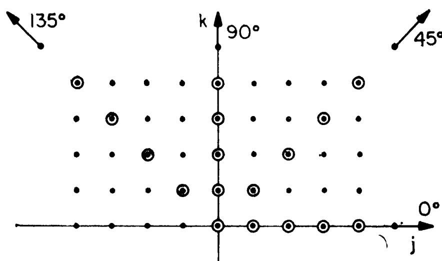


Рис. 17.8.3. Относительное расположение точек при вычислении матриц яркостной зависимости.

$$P(a,\ b;\ j,\ k;\ r,\ \theta) \approx \Pr\{F(j,\ k) = a,\ F(m,\ n) = b\}. \qquad (17.8.4)$$

Для каждого набора параметров $(j,\ k,\ r,\ \theta)$ распределение частот можно рассматривать как массив из $L \times L$ чисел, определяющий степень статистической зависимости пар элементов изображения. Такие массивы называются матрицами яркостной зависимости или матрицами смежности. Распределение, представленное в виде массива из $L \times L$ чисел, необходимо запоминать для каждой точки изображения $(j,\ k)$ и для каждого набора значений $(r,\ \theta)$, поэтому с точки зрения уменьшения вычислительных операций необходимо ограничить угол и модуль вектора расстояния небольшим количеством дискретных значений. На рис. 17.8.3 показано расположение элементов изображения при измерении распределения частот для случая, когда расстояние по радиусу от точки $(j,\ k)$ до точки $(m,\ n)$ принимает четыре дискретных значения, а угол $\theta = 0,\ \pi/4,\ \pi/2,\ 3\pi/4$ радиан (предполагается угловая симметрия). Хорошая оценка совместного распределения вероятностей полу-

чается при достаточно больших частотах. Последнего можно достичь, либо ограничивая число уровней квантования яркости, либо используя сравнительно большие окна. Первый путь ведет к потере

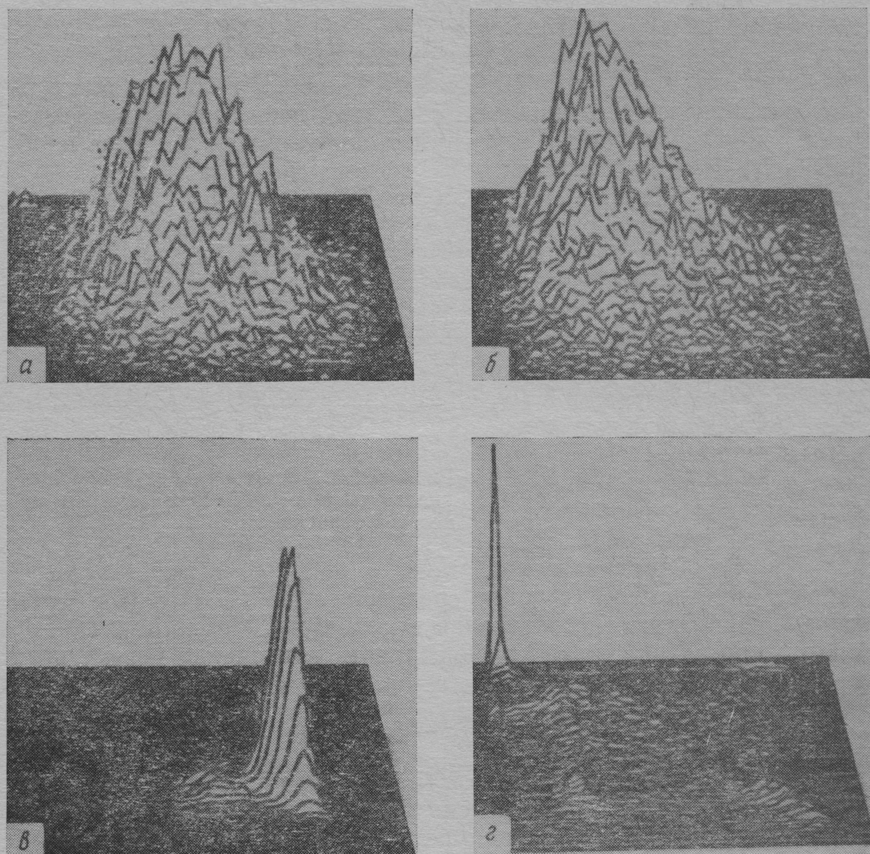


Рис. 17.8.4. Гистограммы, построенные для матриц яркостной зависимости, при $r = 4$, $\theta = 0°$.

а — трава; *б* — плющ; *в* — кирпичная кладка; *г* — решетка

точности при измерении малоконтрастных текстур, а второй дает ошибку, если текстура изменяется в пределах окна. Обычный компромисс состоит в том, чтобы использовать 16 градаций яркости и окно размером примерно 30—50 элементов по каждому измерению.

На рис. 17.8.4 приведены гистограммы яркостей пар элементов для областей с крупной и мелкой текстурой. Для заданных зна-

чений r, θ гистограммы для мелкой текстуры обладают более равномерным распределением по сравнению с гистограммами для крупной текстуры. Размер зерна текстуры может быть измерен величиной рассеяния гистограммы относительно главной диагонали. Харалик и др. [2] предложили ряд мер рассеяния гистограммы для измерения текстуры. Некоторые из них приведены в разд. 17.2. Например, момент инерции, определяемый выражением (17.2.13), дает текстурный признак в виде

$$T(j, k, r, \theta) = \sum_{a=0}^{L-1} \sum_{b=0}^{L-1} (a-b)^2 P(a, b; \; j, k, r, \theta). \quad (17.8.5)$$

Если предположить, что рассматриваемая текстурная область инвариантна относительно поворота, то целесообразно усреднить по углу отдельные текстурные признаки и получить признак текстуры в виде среднего [7]:

$$M_T(j, k, r) = (1/N_\theta) \sum_{\theta} T(j, k, r, \theta), \quad (17.8.6)$$

где суммирование проводится по всем значениям угла, а N_θ — число таких значений. Аналогично дисперсию признака текстуры, не зависящего от угла поворота, можно определить как

$$V_T(j, k, r) = (1/N_\theta) \sum_{\theta} [T(j, k, r, \theta) - M_T(j, k, r)]^2. \quad (17.8.7)$$

Другим полезным параметром является «размах» признака, определяемый как

$$S(j, k, r) = \max_{\theta} \{T(j, k, r, \theta)\} - \min_{\theta} \{T(j, k, r, \theta)\}. \quad (17.8.8)$$

Галлоуэй [37] предложил другой вариант метода измерения текстуры, в котором вместо гистограмм яркостей используются гистограммы длин серий. Длина серии определяется обычным способом как число следующих друг за другом в определенном направлении элементов изображения с одинаковой яркостью. При грубой текстуре получаются длинные серии, а при мелкой — короткие. Несколько признаков текстуры было сформулировано как меры рассеяния гистограммы длин серий.

Наибольшие усилия исследователей в области анализа текстур были направлены не на выявление новых текстурных признаков, а на использование известных признаков для распознавания образов. Например, Харалик и Шанмуган [38] использовали признаки, полученные на основе гистограммы яркостей второго порядка, для классификации спектрозональных аэрофотоснимков, а Уэшка, Дайер и Розенфельд [39] исследовали классификацию местности с помощью нескольких типов текстурных признаков. Крюгер, Томпсон и Тернер [7] применили текстурные признаки для обнаружения и классификации антракоза легких у горняков

по рентгеновским снимкам грудной клетки. Кроме того, Зобрист и Томпсон [40] использовали текстурные признаки для разработки функции, которая оценивает перцептивное различие между текстурными областями.

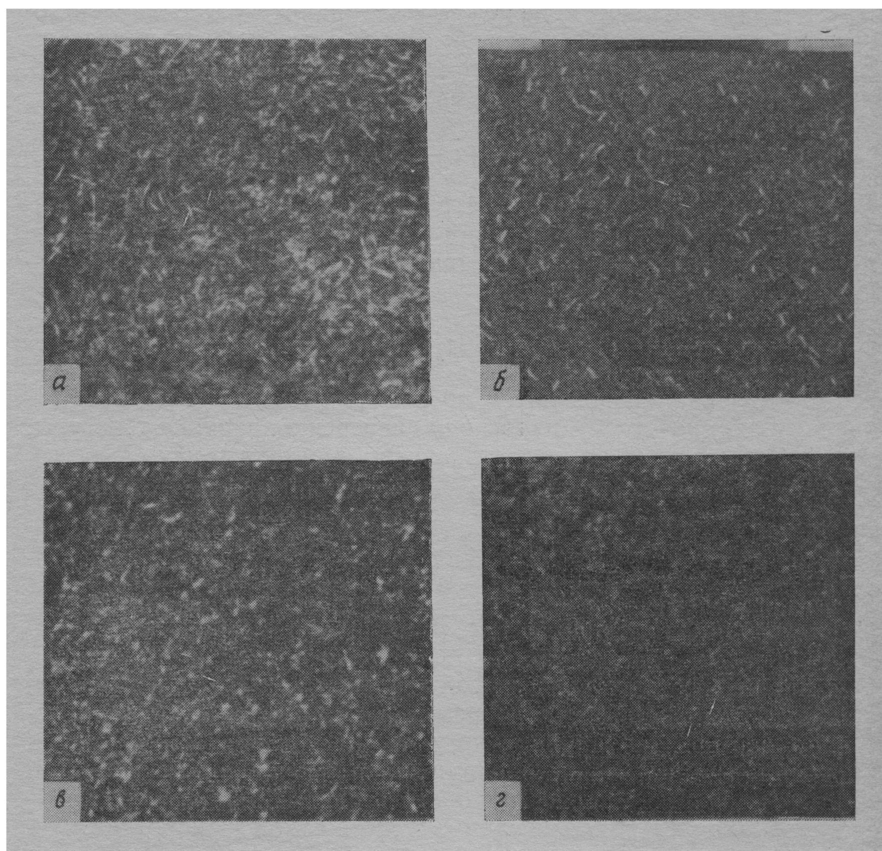


Рис. 17.9.1. Примеры синтеза текстур.
а — естественная трава; *б* — искусственная трава; *в* — естественный плющ; *г* — искусственный плющ.

17.9. СИНТЕЗ ТЕКСТУР

При обработке изображений часто оказывается полезным синтез текстуры. Например, если на картинке недостает какой-либо области или же она сильно искажена шумом, то ее можно заменить областью с искусственной текстурой. Кроме того, при кодировании изображений можно выделить текстурные области, на которых отсутствуют важные детали, и измерить параметры текстуры,

а при декодировании каждую такую область можно заменить искусственной текстурой. По-видимому, такое кодирование окажется более эффективным по сравнению с непосредственным кодированием изображения.

Основной подход к синтезу текстур заключается в создании некоторого базисного первичного образа и в повторении этого образа в плоскости изображения согласно некоторым правилам расстановки [41]. Базисным образом может быть маленький кусочек, выделенный из области с естественной текстурой или же составленный из более мелких элементов, таких, как точка или отрезки прямых линий. Расстановку первичных образов можно выполнять детерминированно, случайным образом или же следовать некоторой смешанной стратегии. Необходимо также учитывать эффекты «ложных» контуров на стыках первичных образов. Обычно необходимо выполнить некоторое сглаживание, чтобы избежать таких эффектов.

Другие способы синтеза текстур основаны на генерации коррелированного случайного поля. Юлес и др. [42] обнаружили, что наблюдатель не может определить разницу между текстурными полями, которые обладают одинаковыми статистиками второго порядка. Так, если имеется соответствующий текстурный признак, то можно построить двумерный процесс второго порядка, соответствующий участку естественной текстуры, при условии, что области с искусственной и естественной текстурами имеют одинаковые значения признака текстуры. На рис. 17.9.1 приведены фотографии естественных текстур — травы, плюща и соответствующих реализаций двумерных случайных полей, имеющих те же статистики второго порядка, что и естественные текстуры.

ЛИТЕРАТУРА

1. Duda R. O., Image Data Extraction (неопубликованные данные), July 1975.
2. Haralick R. M., Shanmugan K., Dinstein I., Texture Features for Image Classification, *IEEE Trans. Systems, Man, and Cybernetics*, **SMC-3**, 610—621 (November 1973).
3. Lendaris G. G., Stanley G. L., An Opticalogical Self-Organizing Recognition System, in: Optical and Electro-Optical Information Processing, Tippett J. T. et al., Eds., Massachusetts Institute of Technology Press, Cambridge, Mass., 1965.
4. Lendaris G. G., Stanley G. L., Diffraction Pattern Sampling for Automatic Pattern Recognition, *Proc. IEEE*, **58**, 2, 198—216 (February 1970). [Имеется перевод: Лендэрис Г. Г., Стенли Г. Л. Метод дискретизации дифракционных картин для автоматического распознавания образов. — ТИИЭР, 1970, т. 58, № 2, с. 22—40.]
5. Haralick R. M., Bosley R., Spectral and Textural Processing of ERTS Imagery, Proceedings 3rd ERTS-1 Symposium, Vol. I, December 1973, pp. 1929—1969.
6. Hall E. L. et al., A Survey of Preprocessing and Feature Extraction Techniques for Radiographic Images, *IEEE Trans. Computers*, **C-20**, 9, 1032—1044 (September 1971).

7. Kruger R. P., Thompson W. B., Turner A. F., Computer Diagnosis of Pneumoconiosis, *IEEE Trans. Systems, Man, and Cybernetics*, **SMC-4**, 1, 40—49 (January 1974).
8. Sutton R. N., Hall E. L., Texture Measures for Automatic Classification of Pulmonary Disease, *IEEE Trans. Computers*, **C-21**, 667—676 (July 1972).
9. Prewitt J. M. S., Object Enhancement and Extraction, in: Picture Processing and Psychopictorics, Lipkin B. S., Rosenfeld A., Eds., Academic Press, New York, 1970.
10. Argyle E., Techniques for Edge Detection, *Proc. IEEE*, **59**, 2, 285—287 (February 1971). [Имеется перевод: Аргайл. Методы обнаружения ступенчатого сигнала. — ТИИЭР, 1971, т. 59, № 1, с. 123—125.]
11. Macleod I. D. G., On Finding Structure in Pictures, in: Picture Language Machines, Kaneff S., Ed., Academic Press, New York, 1970, p. 231.
12. Macleod I. D. G., Comments on Techniques for Edge Detection, *Proc. IEEE*, **60**, 3, 344 (March 1972). [Имеется перевод: Маклеод. Замечания к сообщению «Методы обнаружения ступенчатого сигнала». — ТИИЭР, 1972 т. 60, № 3, с. 126—127.]
13. Rosenfeld A., Thurston M., Lee Y., Edge and Curve Detection: Further Experiments, *IEEE Trans. Computers*, **C-21**, 7, 677—715, (July 1972).
14. Roberts L. G., Machine Perception of Three-Dimensional Solids, in: Optical and Electro-Optical Information Processing, Tippett J. T. et al., Eds., Massachusetts Institute of Technology Press, Cambridge, Mass., 1965, pp. 159—197. [Имеется перевод: Робертс Л. Автоматическое восприятие трехмерных объектов. — В сб. Интегральные роботы, т. 1, М.: Мир, 1973, с. 162—208.]
15. Duda R. O., Hart P. E., Pattern Classification and Scene Analysis, Wiley, New York, 1973. [Имеется перевод: Дуда Р., Харт П., Распознавание образов и анализ сцен. — М.: Мир, 1976.]
16. Kirsch R., Computer Determination of the Constituent Structure of Biological Images, *Computers and Biomedical Research*, **4**, 3, 315—328 (1971).
17. Wallis R. (частное сообщение).
18. Rosenfeld A., A Nonlinear Edge Detection Technique, *Proc. IEEE Letters*, **58**, 5, 814—816 (May 1970). [Имеется перевод: Розенфельд. Нелинейный метод обнаружения ступенчатого сигнала. — ТИИЭР, 1970, т. 58, № 6, с. 94—95.]
19. Rosenfeld A., Thurston M., Edge and Curve Detection for Visual Scene Analysis, *IEEE, Trans. Computers*, **C-20**, 5, 562—569 (May 1971).
20. Hueckel M., An Operator which Locates Edges in Digital Pictures, *JACM*, **18**, 1, 113—125 (January 1971). [Имеется перевод: Хюккель М. Оператор нахождения контуров на кодированных изображениях. — В сб. Интегральные роботы. — М.: Мир, 1973, т. 1, с. 225—240.]
21. Nevatia R., Object Boundary Determination in a Textured Environment, Proceedings ACM'75 Conference, Minneapolis, Minn., October 1975, pp. 32—36.
22. Gouriet G. G., A Method for Measuring Television Picture Detail, *Electronic Engineering*, **24**, 308—311 (July 1952).
23. Kubba M. H., Automatic Picture Detail Detection in the Presence of Random Noise, *Proc. IEEE*, **51**, 11, 1518—1523 (November 1963). [Имеется перевод: Кубба. Автоматическое выделение деталей изображения в присутствии хаотических шумов. — ТИИЭР, 1963, № 11, с. 1491—1497.]
24. Herskovits A., Binford T. O., On Boundary Detection, MIT Project MAC, Artificial Intelligence Memo 183, July 1970.
25. Fram J. R., Deutsch E. S., On the Evaluation of Edge Detection Schemes and Their Comparison with Human Performance, *IEEE Trans. Computers*, **C-24**, 6, 616—628 (June 1975).
26. Robinson G. S., Color Edge Detection, Proceeding SPIE Symposium on Advances in Image Transmission Techniques, Vol. **87**, San Diego, California, August 1976.

27. Nevatia R., Hueckel Color Edge Detector, University of Southern California, Image Processing Institute, Report USCIPI 660, March 1976, pp. 70—81.
28. Cook C. M., Rosenfeld A., Size Detectors, *Proc. IEEE Letters*, **58**, 12, 1956—1957 (December 1970). [Имеется перевод: Кук, Розенфельд. Схема обнаружения объектов заданных размеров. — ТИИЭР, письма, 1970, т. 58, № 12, с. 109, 110.]
29. Zucker S. W., Rosenfeld A , Davis L. S., Picture Segmentation by Texture Discrimination, *IEEE Trans. Computers*, **C-24**, 12, 1228—1233 (December 1975).
30. Heuckel M., A Local Visual Operator Which Recognizes Edges and Lines, *JACM*, **20**, 4, 634—647 (October 1973).
31. Pickett R. M., Visual Analysis of Texture in the Detection and Recognition of Objects, in: Picture Processing and Psychopictorics, Lipkin B. C., Rosenfeld A., Eds., Academic Press, New York, 1970, pp. 289—308.
32. Hawkins J. K., Textural Properties for Pattern Recognition, in: Picture Processing and Psychopictorics, Lipkin B. S., Rosenfeld A., Eds., Academic Press, New York, 1970, pp. 347—370.
33. Brodatz P., Texture: A Photograph Album for Artists and Designers, Dover, New York, 1956.
34. Johnston H. R. et al., Rapid Cartographic Processing System Study, AD 454086, October 1964.
35. Rosenfeld A., Automatic Recognition of Basic Terrain Types from Aerial Photographs, Photogrammic Engineering, **28**, 1, 115—132 (March 1962).
36. Rosenfeld A., Troy E. B., Visual Texture Analysis, Proceedings UMR-Mervin J. Kelly Communications Conference, University of Missouri-Rolla, Missouri, October 1970, Section 10-1.
37. Galloway M. M., Texture Analysis Using Gray Level Run Lengths, *Computer Graphics and Image Processing*, **4**, 2, 172—179 (June 1975).
38. Haralick R. M., Shanmugam K., Computer Classification of Reservoir Sandstones, *IEEE Trans. Geosci. Electronics*, **GE-11**, 171—177 (October 1973).
39. Weska J. S., Dyer C. R., Rosenfeld A., A Comparative Study of Texture Measures for Terrain Classifisation, *IEEE Trans. Systems, Man, and Cybernetics*, **SMC-6**,4, 269—285 (April 1976).
40. Zobrist A. L., Thompson W. B., Building a Distance Function for Gestalt Grouping, *IEEE Trans. Computers*, **C-4**, 7, 718—728 (July 1975).
41. Rosenfeld A, Lipkin B. S., Texture Synthesis, in: Picture Processing and Psychopictorics, Rosenfeld A., Lipkin B. S., Eds., Academic Press, New York, 1970, pp. 309—345.
42. Julesz B. et al., Inability of Humans to Discriminate Between Visual Textures that Agree in Second-Order Statistics-Revisted, *Perception*, **2**, 391—405 (1973).

Глава 18

СИМВОЛИЧЕСКОЕ ОПИСАНИЕ ИЗОБРАЖЕНИЙ

Получение символического описания изображений представляет собой задачу перехода от набора простейших признаков изображения, таких, как значения яркости, контурные точки или параметры текстуры, к значительно меньшему набору средств описания, которые могут служить в качестве исходных данных для последующей семантической интерпретации. Типичными графическими символами являются цепочки контурных точек, образующих границу объекта, связанные области постоянной яркости, цвета или текстуры и элементарные фигуры, такие, как прямоугольники, окружности, треугольники.

В области обработки речевых сигналов существует общепринятый набор речевых символов, названных фонемами, которые являются основными звуками произносимых слов. Например, звук «p» в слове «pin» — фонема. Из фонем теоретически можно сформировать графическое представление устного высказывания. Известен целый ряд разнообразных речевых признаков, включая силу звука, его высоту и энергию спектра. В процессе перехода от тех или иных речевых признаков к фонемам формируется символическое описание речи.

К сожалению, не существует «зрительных фонем», аналогичных фонемам речи, т. е. нет общепринятого набора визуальных символов, которые необходимы и достаточны для описания изображения. Отсутствие набора единых визуальных символов создает определенные трудности при анализе изображений. Во-первых, существует проблема отбора, суть которой состоит в том, чтобы определить, какие символы необходимо сформировать из признаков изображения для решения конкретных задач анализа изображений. Кроме того, имеется проблема определения необходимой точности при формировании символов.

В настоящее время символическое описание изображений находится на самом начальном уровне развития. По-видимому, в лучшем понимании законов построения символического описания лежит ключ к возможному развитию общих систем анализа изображений. В настоящей главе дается обзор способов формирования символов изображения.

18.1. СВЯЗНОСТЬ

Основной этап при формировании символического описания изображения по массиву элементов или набору простейших признаков заключается в определении геометрических соотношений и связности между элементами, относительно которых предполагается, что они принадлежат одному классу [1, 2]. Для двоичной картинки, представленной на рис. 18.1.1, *а*, кольцо из четырех элементов, согласно всем принятым определениям связности, делит эту картинку на три области: белые элементы с внешней стороны кольца, белые элементы внутри кольца и черные элементы самого кольца. Говорят, что элементы внутри каждой области связаны друг с другом. Смысл этого понятия легко уяснить, обратившись к рис. 18.1.1, *а*, но если рассматривать рис. 18.1.1, *б*, то возникает неоднозначность. Определяют ли все черные элементы кольцо или же они представляют собой четыре прямолинейных отрезка? Ответ на этот вопрос до некоторой степени зависит от желаемого определения связности.

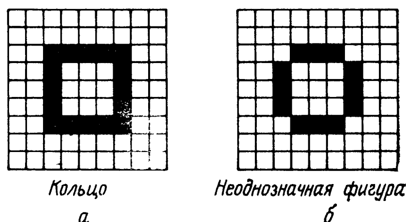


Рис. 18.1.1. К определению связности. *а* — кольцо; *б* — неоднозначная фигура.

Возвращаясь к более общему случаю многоградационных изображений, рассмотрим рис. 18.1.2, *а*, на котором представлен некоторый элемент (элемент *А*), окруженный восемью соседними элементами (от *В* до *I*). Предположим, что элемент *А* обладает свойством *S*, установленным на основе некоторого простейшего описания (яркости, цвета, текстуры и т. д.). По определению четырехсвязности [1]) элемент *А* и элемент *В* связаны, если оба обладают свойством *S*. Аналогично четырехсвязность можно установить между элементом *А* и элементами *Е*, *С* и *D*, граничащими с *А* по ребру, при условии, что оба члена пары обладают одним и тем же свойством. Восьмисвязность позволяет связывать элемент *А* с одним из его соседей по диагонали, например с элементом *F*, граничащим с *А* в точке, если оба они обладают одинаковым свойством.

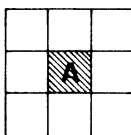
На рис. 18.1.1, *б* в соответствии с определением четырехсвязности имеется четыре несвязных черных прямолинейных отрезка, а согласно определению восьмисвязности, изображено кольцо из связных черных элементов. Заметим, однако что при восьмисвязности белые элементы, расположенные внутри кольца на

[1]) Под четырехсвязностью здесь понимается связность по четырем направлениям: вверх, вниз и влево, вправо. — *Прим. перев.*

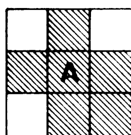
рис. 18.1.1, *б*, связаны с белыми элементами с внешней стороны кольца. Таким образом возникает парадокс. Если бы черные элементы связывались по принципу восьмисвязности в кольцо, то следовало бы ожидать разделения внутренних и внешних белых элементов этого кольца. Чтобы разрешить эту дилемму, можно для элементов со свойством S определить восьмисвязность, а принцип четырехсвязности установить для элементов, обладающих свойством $\bar{S}$ ($\bar{S}$ — дополнение множества S), или наоборот.

G	C	F
D	A	B
H	E	I

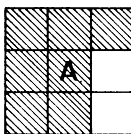
а



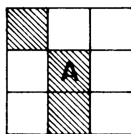
б



в



г



д

Рис. 18.1.2. Определения связных элементов изображения.

а — обозначения элементов окрестности; *б* — изолированный элемент; *в* — внутренний элемент; *г* — граничный элемент; *д* — элемент дуги.

Обратимся к рис. 18.1.2. Пусть заштрихованный элемент обладает свойством S, а незаштрихованный — свойством $\bar{S}$. Тогда элемент A на рис. 18.1.2, *б* называется изолированным, если для него не соблюдается принцип восьмисвязности относительно любого из его соседей. На рис. 18.1.2, *в* элемент A является внутренним элементом, для которого выполняется принцип четырехсвязности относительно каждого из его соседей B, C, D, E. Граничный элемент, как показано на рис. 18.1.2, *г*, не обладает четырехсвязностью по крайней мере с одним из ближайших соседей. Следуя этому определению, элемент C из окрестности A не может быть классифицирован как граничная точка. Рис. 18.1.2, *д* иллюстрирует определение точки дуги; элемент A обладает четырехсвязностью только со своими верхним и нижним (или правым и левым) соседями. Дуговой концевой элемент обладает четырехсвязностью лишь с одним соседом. Наконец, минимально связная дуга по определению есть множество точек дуги, для которых каждая внутренняя точка дуги (не являющаяся ее концом) обладает восьмисвязностью лишь с двумя соседями.

Предыдущие определения связности основаны на цифровой модели изображения, полученной посредством взятия отсчетов на поле непрерывного изображения в точках прямоугольного растра.

Голэй [3] предложил гексагональную решетчатую структуру, показанную на рис. 18.1.3. Такая структура позволяет обойти многие трудности, обусловленные применением прямоугольной решетки. Говорят, что в случае гексагональной системы соседние элементы обладают шестисвязностью, если они имеют одинаковые свойства и граничат друг с другом по общему ребру. Для многих задач выделения признаков разработаны алгоритмы построения границ областей, использующие шестисвязность [4]. Однако широкому использованию гексагональной системы препятствуют два основных ее недостатка. Во-первых, большинство устройств развертки изображений имеет прямоугольный растр; переход к гексагональному растру сопряжен с большими трудностями. Во-вторых, гексагональный растр не очень удобен для многих алгоритмов пространственной обработки, например таких, как свертка и фильтрация с преобразованием Фурье.

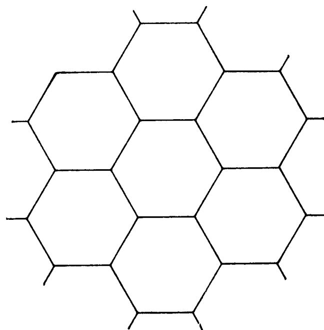


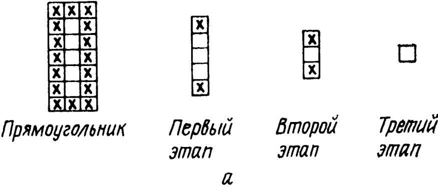
Рис. 18.1.3. Гексагональная решетка элементов изображения.

18.2. СЖАТИЕ, УТОНЧЕНИЕ И ПОСТРОЕНИЕ ОСТОВА

Сжатие и утончение представляют собой необратимые операции, цель которых состоит в том, чтобы попытаться свести связные области элементов изображения с заданным набором свойств к областям меньших размеров. Посредством сжатия область сводится к единственному элементу, а с помощью утончения область приводится к минимальной ширине поперечного сечения. Построение остова — это операция, с помощью которой область преобразуется в фигуру, напоминающую каркас.

На рис. 18.2.1 приведены примеры, иллюстрирующие обработку объекта прямоугольной формы и области неправильной формы простым алгоритмом сжатия. С помощью этого алгоритма граничные точки, не являющиеся точками дуги (отмеченные на рисунке знаком ×), удаляются из области, если удаление элемента не ведет к нарушению связности области согласно определению восьмисвязности. Точки дуги удаляются лишь в том случае, если они являются концевыми точками дуги и их удаление не приводит к исчезновению области. Алгоритм заканчивает свою работу, когда остается единственный элемент.

Работа простого алгоритма утончения иллюстрируется на рис. 18.2.2, *а* на примере утончения объекта прямоугольной формы. На первом шаге первого этапа работы алгоритма граничные эле-



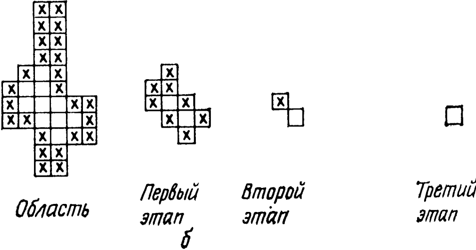
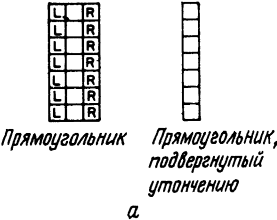


Рис. 18.2.1. Примеры работы алгоритма сжатия.
а — сжатие объекта прямоугольной формы; *б* — сжатие области.



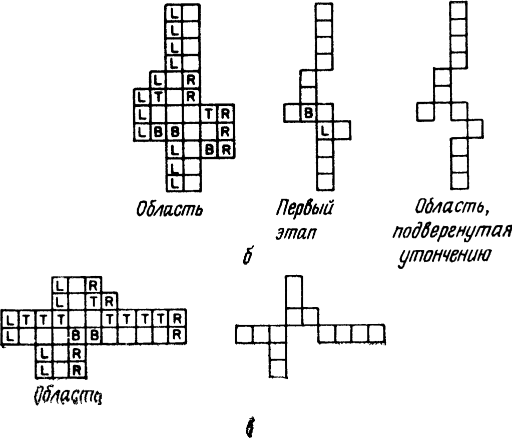


Рис. 18.2.2. Примеры работы алгоритма утончения.
а — утончение прямоугольника; *б* — утончение вертикально ориентированной области; *в* — утончение горизонтально ориентированной области.

менты с левой стороны объекта, обозначенные буквой L, удаляются, если они не являются точками дуги и их удаление не ведет к нарушению восьмисвязности. На втором шаге удаляются граничные элементы с правой стороны объекта, обозначенные буквой R, если соблюдаются такие же условия, как и для левых граничных точек. Затем процесс удаления повторяется для верхних (T) и нижних (B) граничных точек, которые удаляются, если они не являются точками дуги и их удаление не ведет к нарушению восьмисвязности. После четырех шагов первого этапа работа алгоритма повторяется до тех пор, пока нельзя будет удалить ни один элемент без нарушения связности. На рис. 18.2.2, *б* и *в* приведены примеры, иллюстрирующие работу этого алгоритма применительно к области неправильной формы, ориентированной вертикально и горизонтально. Результаты получаются различные вследствие заданной последовательности выполнения шагов.

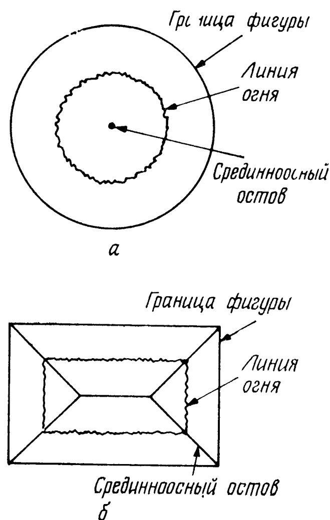


Рис. 18.2.3. Примеры преобразований к срединным осям.
а — круг; *б* — прямоугольник.

Розенфельд [1, 2, стр. 82] и Престон [5] разработали и проанализировали несколько последовательных алгоритмов сжатия и утончения. К настоящему времени разработаны также алгоритмы для нескольких машин параллельной обработки [6—9].

Чтобы выразить структурные соотношения сложных объектов в сцене, часто оказывается достаточным представление объектов в виде остова, или каркаса. Ясно, что остов большинства объектов, как правило, можно представить значительно эффективнее, чем сам объект. Один из подходов к получению остова заключается в утончении объекта до тех пор, пока не будет получена цепочка элементов с минимальной связностью. Недостаток этого подхода состоит в том, что остов определяется неоднозначно; форма получающейся в результате фигуры обычно сильно зависит от алгоритма утончения.

Блюм [10] предложил остроумный способ получения остова, названный преобразованием к срединным осям, который для каждого заданного объекта дает ознозначный результат. Интуитивное определение такого преобразования основывается на аналогии со «степным пожаром» [11—13]. Рассмотрим изображенные на рис. 18.2.3 области круглой и прямоугольной формы и представим себе, что это участки земли, покрытые высохшей травой. Если бы

огонь возник одновременно по всему периметру участков, то он распространялся бы к их центрам до тех пор, пока не сгорела бы вся трава. В случае круглой области огонь распространялся бы к центру круга, который представляет собой точку самогашения огня. Для прямоугольной области огонь распространялся бы с каждой стороны. По мере продвижения огня слева и сверху линии огня будут встречаться и пожар будет затухать. Геометрическое место точек самогашения огня образует линию самогашения. Точки или линии самогашения называются срединными осями или остовом фигуры. Вообще срединноосный остов состоит из множества точек, которые находятся на одинаковом расстоянии от двух ближайших к ним точек на границе фигуры. Это расстояние называется расстоянием самогашения фигуры. Границу фигуры можно восстановить по ее срединноосному остову и расстоянию самогашения. Эта граница есть огибающая окружностей с радиусом, равным расстоянию самогашения, с центрами в каждой точке срединноосного остова.

Применение на практике преобразования к срединным осям дискретных объектов, определенных на прямоугольном растре, связано с определенными трудностями. Чтобы обеспечить однозначность преобразования, необходимо надлежащим образом ввести расстояние. Если объект имеет границу сложной формы, его остов также будет сложным. С целью достижения более компактного представления фигуры Монтанари [14, 15] исследовал влияние упрощения остова объекта. Для многих объектов стирание или сглаживание частей остова приводит к получению приблизительной границы объекта, которая не отличается заметно от истинной.

18.3. ОПИСАНИЕ ЛИНИЙ

Прямолинейные и криволинейные отрезки образуют основную структуру многих изображений. Для таких изображений математические соотношения между выделенными точками на границе объекта дают символическое описание изображения. В настоящем разделе описываются два подхода к установлению математических соотношений: подбор кривых и преобразование линии в точку.

18.3.1. АППРОКСИМАЦИЯ КРИВЫХ

Рассмотрим совокупность точек (x_i, y_i) при $i = 0, 1, 2, \ldots, M$, взятых на границе двумерного объекта, как показано на рис. 18.3.1. Предположим, что эти точки упорядочены в том смысле, что точки (x_i, y_i) и (x_{i+1}, y_{i+1}) — ближайшие соседи вдоль границы. Если $x_{i+1} > x_i$, то говорят, что эти точки находятся в функциональной связи. Типичным примером функционально связанных точек служат точки графика зависимости напряжения от времени в электрической цепи.

Аппроксимация кривой на множестве точек состоит в определении некоторой функции $\hat{y} = g(x)$, для которой ошибка аппроксимации, т. е. мера отклонения совокупности исходных точек (x_i, y_i) от точек $[x_i, g(x_i)]$, принимает минимальное значение. Если точки объекта находятся в функциональной связи, то ошибка аппроксимации обычно измеряется по координате y. Типичными ошибками аппроксимации являются

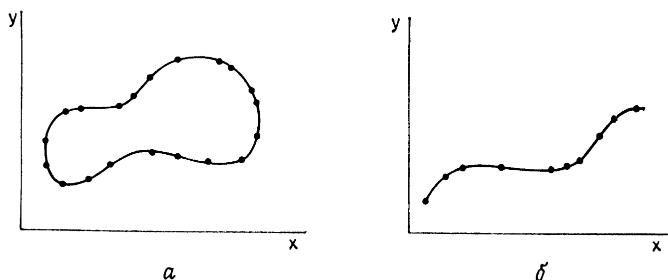


Рис. 18.3.1. Точки, взятые на границе объекта.
а — точки объекта; *б* — точки, находящиеся в функциональной связи.

абсолютная ошибка

$$\mathscr{E} = \sum_{i=1}^{M} |y_i - g(x_i)|, \tag{18.3.1а}$$

квадратическая ошибка

$$\mathscr{E} = \sum_{i=1}^{M} [y_i - g(x_i)]^2, \tag{18.3.1б}$$

максимальная ошибка

$$\mathscr{E} = \max_{i} |y_i - g(x_i)|. \tag{18.3.1в}$$

Для общего случая произвольно заданных точек объекта меры ошибки, заданные выражениями (18.3.1), часто оказываются бессмысленными. При этом ошибку, которая связана с каждой точкой (x_i, y_i), можно измерить расстоянием от этой точки до аппроксимирующей кривой по нормали к ней и аналогично формулам (18.3.1) записать выражения для абсолютной квадратической и максимальной ошибок. Минимизация получающейся в результате ошибки аппроксимации для общего случая обычно представляет собой довольно трудную задачу.

Среди наиболее часто используемых способов аппроксимации кривых для функционально связанных экспериментальных точек

следует указать кусочно-полиномиальную аппроксимацию. В этом случае аппроксимирующая функция имеет вид

$$\hat{y} = a_0 + a_1 x + a_2 x^2 + \cdots + a_N x^N, \tag{18.3.2}$$

где a_n — коэффициенты полинома. Подстановка экспериментальных точек в выражение (18.3.2) приводит к соотношению в векторной форме

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^N \\ 1 & x_1 & x_1^2 & \ldots & x_1^N \\ 1 & x_2 & x_2^2 & \ldots & x_2^N \\ & & \cdot & & \\ & & \cdot & & \\ & & \cdot & & \\ 1 & x_M & x_M^2 & \ldots & x_M^N \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_N \end{bmatrix} = \begin{bmatrix} \hat{y}_0 \\ \hat{y}_1 \\ \hat{y}_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{y}_M \end{bmatrix}, \tag{18.3.3а}$$

которое можно записать в компактном виде

$$\mathbf{Xa} = \hat{\mathbf{y}}. \tag{18.3.3б}$$

Для ошибки по критерию наименьших квадратов

$$\mathscr{E} = (\mathbf{y} - \hat{\mathbf{y}})^T (\mathbf{y} - \hat{\mathbf{y}}) \tag{18.3.4}$$

оптимальный набор коэффициентов полинома получается из уравнения

$$\mathbf{a} = \mathbf{X}^{-}\mathbf{y}. \tag{18.3.5}$$

При $M > N$, т. е. когда число экспериментальных точек превышает число коэффициентов полинома, псевдообратная матрица вычисляется по формуле

$$\mathbf{X}^{-} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T \tag{18.3.6}$$

при условии, что все экспериментальные точки x_i различны. В случае линейной аппроксимации требуется определить лишь коэффициенты a_0 и a_1. В противоположном крайнем случае, когда функция $\hat{y} = g(x)$ представляет собой полином n-го порядка, имеет место равенство $\mathbf{X}^{-} = \mathbf{X}^{-1}$ и уравнение (18.3.3) справедливо для каждой экспериментальной точки, т. е. аппроксимирующий полином проходит через каждую экспериментальную точку. В этом случае интерполирующий полином единственный, но его можно представить и подсчитать различным образом, например по формулам интерполяции Лагранжа, Ньютона, Эйткена [16].

В работе Дуда и Харта [11] отдается должное Форсену как родоначальнику простого метода кусочно-линейной аппроксимации, названного итеративным подбором концевых точек. На первом

этапе работы алгоритма, которая иллюстрируется рис. 18.3.2, концевые экспериментальные точки A и B соединяются прямой линией. Затем исследуется точка с наибольшим отклонением от этой прямой (точка C). Если отклонение достаточно велико, то эта точка берется в качестве точки соединения двух отрезков (AC и CB). Данная процедура повторяется для каждого отрезка до тех пор, пока экспериментальные точки не будут «хорошо» аппроксимированы отрезками прямых. Основное преимущество этого алгоритма состоит в его простоте, а недостаток — в возникновении ошибок, обусловленных ошибками в исходных экспериментальных данных. Реймер [17] использовал способ, подобный процедуре итеративного подбора концевых точек, для полигональной аппроксимации замкнутых кривых произвольной формы. Павлидис и Хоровиц [18] разработали алгоритм для полигональной аппроксимации кривых.

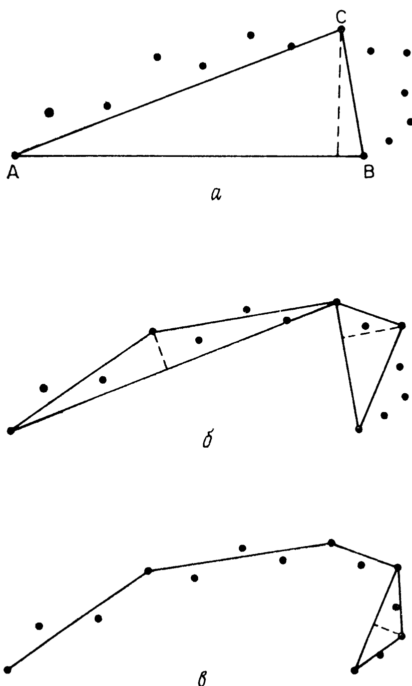


Рис. 18.3.2. Итеративный подбор концевых точек.

а — первый этап; *б* — второй этап; *в* — третий этап.

18.3.2. ПРЕОБРАЗОВАНИЕ ЛИНИИ В ТОЧКУ

Другой подход к описанию линии, развитый Хохом [19], предусматривает преобразование линии, заданной в декартовой системе координат, в точку в полярной системе координат. Прямую линию (рис. 18.3.3, *а*) можно описать в параметрической форме в виде уравнения

$$\rho = x \cos \theta + y \sin \theta, \tag{18.3.7}$$

где ρ — расстояние, измеренное от начала координат по нормали к прямой, а θ — угол наклона этой нормали относительно оси x. Преобразование Хоха, примененное к прямой линии, дает точку (ρ, θ) на плоскости с полярными координатами (рис. 18.3.3, *б*). Семейство прямых, проходящих через общую точку (рис. 18.3.3, *в*), отображается в совокупность точек в системе координат ρ, θ (рис. 18.3.3, *г*). Теперь рассмотрим три коллинеарные точки

(рис. 18.3.3, *д*). Преобразование Хоха, примененное к семействам прямых, проходящих через эти точки, дает в результате совокупность трех параметрических кривых в системе ρ, θ (рис. 18.3.3, *е*).

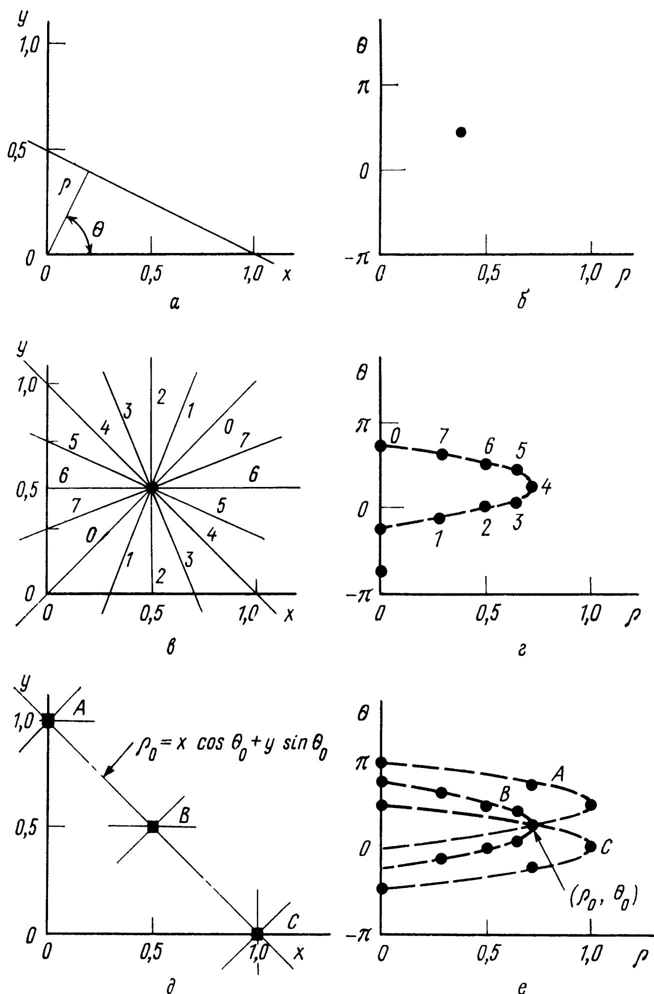


Рис. 18.3.3. Примеры преобразования Хоха.

а — параметрическое задание прямой; *б* — преобразование Хоха прямой *а*; *в* — семейство прямых, проходящих через общую точку; *г* — преобразование Хоха семейства прямых *в*; *д* — коллинеарные точки; *е* — преобразование Хоха точек *д*.

Эти три кривые пересекаются в единственной точке с координатами ρ_0, θ_0, которая соответствует пунктирной прямой, проходящей на рис. 18.3.3, *д* через три коллинеарные точки.

Дуда и Харт [20] использовали преобразование Хоха для обнаружения прямых и кривых линий на цифровых изображениях. Каждая дискретная наблюдаемая точка на плоскости x, y преобразуется в кривую на плоскости ρ, θ [1]). Эта плоскость разбивается на ячейки. Если какой-либо участок кривой проходит через некоторую ячейку, то вес данной ячейки возрастает на единицу. После того как все наблюдаемые точки будут подвергнуты преобразованию, исследуются ячейки плоскости ρ, θ. Большой вес ячейки соответствует коллинеарным экспериментальным точкам, которые можно аппроксимировать прямой линией с параметрами, приблизительно равными ρ, θ. Малый вес ячейки обычно указывает на изолированную точку, которую можно исключить из рассмотрения. О'Горман и Клоуз [21] предложили видоизмененное преобразование Хоха для связывания точек перепадов яркости в контурную линию. В их процедуре угол θ на плоскости ρ, θ получается исходя из направления градиента перепада. Затем по формуле (18.3.7) подсчитывается соответствующее значение ρ для координат перепада x, y. Однако вес ячейки (ρ, θ) увеличивается не на единицу, а на величину модуля градиента с тем, чтобы придать большую значимость сильным (контрастным) перепадам по сравнению со слабыми.

18.4. ОПИСАНИЕ ФОРМЫ

Прямые и кривые линии можно рассматривать как простейшие элементы более крупных структур, например таких, как прямоугольники, треугольники, окружности и пятна произвольной формы. Эти структуры затем можно описать и проанализировать с помощью характеристик их формы: метрических, топологических и аналитических. В данном разделе рассматривается только машинный анализ формы. В работах [22—24] описываются исследования формы с точки зрения человеческого восприятия.

18.4.1. МЕТРИЧЕСКИЕ ХАРАКТЕРИСТИКИ

Метрические характеристики изображения основаны на измерении расстояния между точками на его плоскости [11, 25]. Расстояние — это вещественная функция $d\,[(x_i, y_i), (x_j, y_j)]$ двух точек (x_i, y_i) и (x_j, y_j), обладающая следующими свойствами:

$$d\,[(x_i, y_i), (x_j, y_j)] \geqslant 0, \tag{18.4.1а}$$

$$d\,[(x_i, y_i), (x_j, y_j)] = d\,[(x_j, y_j), (x_i, y_i)], \tag{18.4.1б}$$

$$d\,[(x_i, y_i), (x_j, y_j)] + d\,[(x_j, y_j), (x_k, y_k)] \geqslant \geqslant d\,[(x_i, y_i), (x_k, y_k)]. \tag{18.4.1в}$$

[1]) См. рис. 18.3.3, *в, г.* — *Прим. ред.*

Известно довольно много функций, удовлетворяющих условиям (18.4.1). Большинство обычных метрик, встречающихся в задачах анализа изображений, имеет следующий вид:

евклидово расстояние

$$d_E = [(x_i - x_j)^2 + (y_i - y_j)^2]^{1/2}, \quad (18.4.2а)$$

абсолютное расстояние

$$d_M = |x_i - x_j| + |y_i - y_j|, \quad (18.4.2б)$$

максимальное расстояние

$$d_X = \max\{|x_i - x_j|, |y_i - y_j|\}. \quad (18.4.2в)$$

В случае дискретного изображения разности координат $x_i - x_j$ и $y_i - y_j$ представляют собой целые числа, а евклидово расстояние обычно не целочисленно. Это обстоятельство неизбежно ведет к ошибке вследствие округления или усечения числа при цифровой обработке.

Установив метрику, можно разработать различные метрические характеристики изображения. Среди наиболее важных — длина периметра P и площадь объекта A. Другой полезной характеристикой является коэффициент, вычисляемый по формуле

$$T = 4\pi (A/P^2). \quad (18.4.3)$$

18.4.2. ТОПОЛОГИЧЕСКИЕ ХАРАКТЕРИСТИКИ

Топологические характеристики формы — это свойства, инвариантные по отношению к преобразованию «резинового листа» [11, 26—28]. Такое преобразование или отображение можно представить себе как растяжение резинового листа с нарисованным объектом заданной формы, в результате которого происходит некоторое пространственное искажение фигуры. Преобразования, требующие разрывов резинового листа или соединения одной его части с другой, не допустимы. Ясно, что расстояние не является топологическим свойством, поскольку оно может изменяться при растяжении резинового листа. Такие понятия, как перпендикулярность и параллельность линий, также не являются топологическими свойствами. На рис. 18.4.1, *а* представлена двоичная фигура, состоящая из двух связных компонент, а на рис. 18.4.1, *б* изображена та же фигура, подвергнутая растяжению. Понятно, что нет таких операций растяжения, которые смогли бы увеличить или уменьшить связность фигуры. Связные компоненты изображения могут содержать дыры (рис. 18.4.1, *в*). Количество дыр, очевидно, не изменяется при топологическом отображении. Фунда-

ментальное соотношение между количеством связных компонент C и числом дыр H на фигуре, называемое числом Эйлера, имеет вид

$$E = C - H. \tag{18.4.4}$$

Число Эйлера — также топологическое свойство, поскольку C и H — топологические свойства.

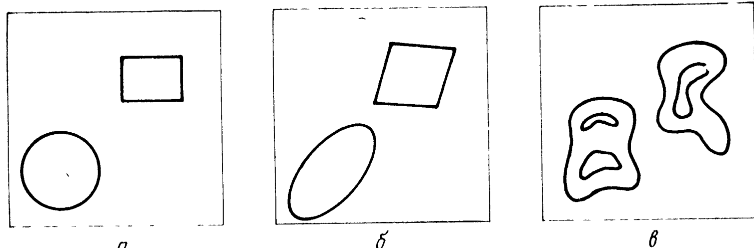


Рис. 18.4.1. Примеры топологических свойств.

а — объекты, $C = 2$, *б* — объекты на резиновом листе, подвергнутом растяжению; *в* — объекты с дырами, $C = 2$, $H = 3$, $E = -1$.

Объекты неправильной формы можно описать как совокупность их топологических составляющих. Рассмотрим трубчатый объект, сечение которого напоминает букву *R*, изображенную на рис. 18.4.2, *а*, и представим, что на объект натянута резиновая лента. Область, ограниченная резиновой лентой, называется выпуклой оболочкой объекта. Множество точек внутри выпуклой оболочки, не принадлежащих объекту, образует дефицит выпуклости объекта. Имеются два типа дефицита выпуклости: области, полностью ограниченные объектом, которые названы озерами, и области, лежащие между периметром выпуклой оболочки и объектом, которые названы заливами. Для некоторых применений проще описать объект непосредственно с использованием понятий выпуклой оболочки и дефицита выпуклости. Для объектов, представленных на прямолинейных решетках, определение выпуклой оболочки необходимо несколько видоизменить с тем, чтобы оставался прежний смысл. Такие объекты, как окружности и треугольники, на дискретном растре следует рассматривать как выпуклые, даже если их границы имеют ступенчатый вид. Эту кажущуюся трудность

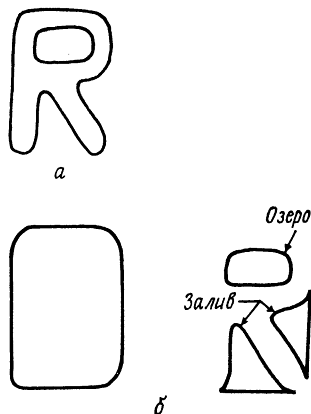


Рис. 18.4.2. Средства описания выпуклой формы.

а — фигура; *б* — выпуклая оболочка, заливы и озера.

можно обойти, рассматривая резиновую ленту, натянутую на дискретизованный объект. Элементы, полностью лежащие внутри объема, ограниченного резиновой лентой, и не принадлежащие объекту, образуют дефицит выпуклости. Скланский и др. [29, 30] разработали практические алгоритмы для вычисления характеристик выпуклости дискретизованных объектов.

Топологические свойства, определенные выше, пригодны для использования в качестве символов при распознавании объектов на изображениях. Кроме того, топологические свойства можно применять для эффективного вычисления геометрических величин, например таких, как периметр и площадь, используя методы, основанные на нахождении участков изображения, совпадающих с тем или иным из заданных двоичных эталонов [28]. Рассмотрим объект, определенный на прямоугольной решетке, элементы которого имеют значение единица, а элементы фона — значение нуль. Площадь объекта равна, очевидно, числу единиц в массиве значений элементов. Символически это обозначается так:

$$A = n\,(1), \tag{18.4.5}$$

где $n\,(\cdot)$ — число совпадений с указанным в скобках эталоном. Если объект полностью окружен белыми элементами, то его периметр равен

$$P = 2\cdot n\,(0\;1) + 2\cdot n\begin{pmatrix}0\\1\end{pmatrix}. \tag{18.4.6}$$

Например, для квадрата размером 2×2 элементов $A = 4$ и $P = 8$. Объект, образованный из трех диагонально связанных элементов, имеет $A = 3$ и $P = 12$. Рассмотрим теперь следующий набор эталонов размером 2×2 элементов, названных двоичными четверками:

$$Q_0 = \left\{\begin{pmatrix}0 & 0\\0 & 0\end{pmatrix}\right\}, \tag{18.4.7а}$$

$$Q_1 = \left\{\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}\right\}, \tag{18.4.7б}$$

$$Q_2 = \left\{\begin{pmatrix}1 & 1\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\0 & 1\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\1 & 0\end{pmatrix}\right\}, \tag{18.4.7в}$$

$$Q_3 = \left\{\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\1 & 1\end{pmatrix}, \begin{pmatrix}1 & 1\\1 & 0\end{pmatrix}\right\}, \tag{18.4.7г}$$

$$Q_4 = \left\{\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}\right\}, \tag{18.4.7д}$$

$$Q_D = \left\{\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\right\}. \tag{18.4.7е}$$

Площадь и периметр изображения можно связать с количеством входящих в него двоичных четверок формулами

$$A = {}^1/_4\,[n(Q_1) + 2\cdot n(Q_2) + 3\cdot n(Q_3) + 4\cdot n(Q_4) + 2\cdot n(Q_D)], \quad (18.4.8а)$$

$$P = n(Q_1) + n(Q_2) + n(Q_3) + 2\cdot n(Q_D). \quad (18.4.8б)$$

Эти формулы могут давать значительную ошибку, если вычисленные по ним значения периметра и площади использовать как оценки этих параметров для непрерывного объекта, который был подвергнут дискретизации. Более точные формулы для таких оценок были получены Дудой [31]:

$$A = ({}^1/_4)\,n(Q_1) + ({}^1/_2)\,n(Q_2) + ({}^7/_8)\,n(Q_3) + n(Q_4) + ({}^3/_4)\,n(Q_D), \quad (18.4.9а)$$

$$P = n(Q_2) + \left(1/\sqrt{2}\right)[n(Q_1) + n(Q_3) + 2\cdot n(Q_D)]. \quad (18.4.9б)$$

Подсчет двоичных четверок позволяет очень просто определить число Эйлера изображения. Грэй [28] установил, что в соответствии с определением четырехсвязности число Эйлера можно вычислить следующим образом:

$$E = ({}^1/_4)\,[n(Q_1) - n(Q_3) + 2\cdot n(Q_D)], \quad (18.4.10а)$$

а для восьмисвязности

$$E = ({}^1/_4)\,[n(Q_1) - n(Q_3) - 2\cdot n(Q_D)]. \quad (18.4.10б)$$

Если на изображении имеется много связных компонент и мало дыр, то число Эйлера можно принять в качестве оценки числа компонент. Тогда среднюю площадь и средний периметр связных компонент можно выразить как [28]

$$\overline{A} \approx A/E, \quad (18.4.11а)$$

$$\overline{P} = P/E. \quad (18.4.11б)$$

Для изображений, содержащих тонкие объекты, например машинописные или рукописные буквы, приближенные значения средней длины и средней ширины объекта можно вычислить по формулам [28]

$$\overline{L} = \overline{P}/2, \quad (18.4.12а)$$

$$\overline{W} = 2\overline{A}/\overline{P}. \quad (18.4.12б)$$

Эти простые меры оказываются полезными для построения грубых различительных характеристик изображения, указывающих, со-

держит ли изображение много маленьких точечных объектов или же небольшое число пятен большого размера, имеются ли толстые или тонкие объекты и т. д.

18.4.3. АНАЛИТИЧЕСКИЕ ХАРАКТЕРИСТИКИ

Аналитические характеристики формы — это математические описания, которые дают несколько иное ее представление. Конечно, чтобы эти описания оказались полезными, они должны быть проще, чем исходное представление формы в виде массива значений дискретных отчетов.

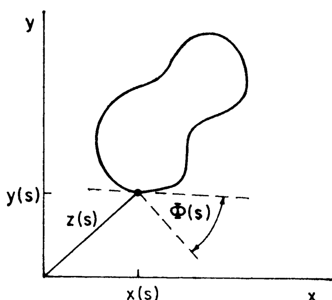


Рис. 18.4.3. Определение кривизны.

Периметр произвольной замкнутой кривой можно представить через совокупность текущих значений кривизны в каждой его точке. Рассмотрим непрерывную замкнутую кривую, вычерченную в комплексной плоскости на рис. 18.4.3, для которой положение точки на периметре задается ее координатой $z(s)$, зависящей от длины s. Комплексную функцию $z(s)$ можно выразить в виде ее действительной $x(s)$ и мнимой $y(s)$ составляющих:

$$z(s) = x(s) + iy(s). \tag{18.4.13}$$

Угол наклона касательной, показанный на рис. 18.4.3, определяется выражением

$$\Phi(s) = \operatorname{arctg}\left[\frac{dy(s)/ds}{dx(s)/ds}\right], \tag{18.4.14}$$

а кривизна представляет собой действительную функцию

$$k(s) = d\Phi(s)/ds. \tag{18.4.15}$$

Зная функцию кривизны, координаты точек $x(s)$, $y(s)$ можно получить по формулам

$$x(s) = x(0) + \int_0^s k(\alpha)\cos[\Phi(\alpha)]\,d\alpha, \tag{18.4.16а}$$

$$y(s) = y(0) + \int_0^s k(\alpha)\sin[\Phi(\alpha)]\,d\alpha, \tag{18.4.16б}$$

где $x(0)$ и $y(0)$ — координаты начальной точки.

Поскольку функция кривизны периодическая с периодом, равным длине периметра P, ее можно разложить в ряд Фурье

$$k(s)=\sum_{n=-\infty}^{\infty} c_n \exp\left\{\frac{2\pi i}{P} ns\right\} \qquad (18.4.17а)$$

с коэффициентами Фурье

$$c_n=\frac{1}{P}\int_0^P k(s)\exp\left\{-\frac{2\pi i}{P} n\right\} ds. \qquad (18.4.17б)$$

На этой основе Косгриф [32], Брилл [33] и другие исследователи разработали способ получения набора фурье-описаний с помощью разложения функции формы в ряд Фурье при ограничении несколькими членами ряда. Эти описания используются для символического представления формы при последующем распознавании и анализе.

Если форма имеет резкие изломы, как, например, прямоугольник, функция кривизны в точках разрыва оказывается неопределенной. Эту аналитическую трудность можно обойти, используя функцию формы

$$\theta(s)=\int_0^s k(\alpha)\,d\alpha-\frac{2\pi s}{P}, \qquad (18.4.18)$$

предложенную Цанем [34]. Эта функция также периодическая с периодом P, и, следовательно, для получения описания формы ее можно раскладывать в ряд Фурье.

Беннетт и Макдональд [35] проанализировали ошибку дискретизации, вызванную тем, что функция кривизны определяется на дискретном изображении. Для дискретного случая кривизна определяется формулами

$$z(s_j)=x(s_j)+iy(s_j), \qquad (18.4.19а)$$

$$\Phi(s_j)=\operatorname{arctg}\left[\frac{y(s_j)-y(s_{j-1})}{x(s_j)-x(s_{j-1})}\right], \qquad (18.4.19б)$$

$$k(s_j)=\Phi(s_j)-\Phi(s_{j-1}), \qquad (18.4.19в)$$

где s_j — j-й элемент дуги. На рис. 18.4.4 приведены результаты фурье-разложения дискретной функции кривизны при использовании простого правила четырехсвязности.

Другой подход к получению аналитического описания формы состоит в аппроксимации посредством моментов. В теории вероят-

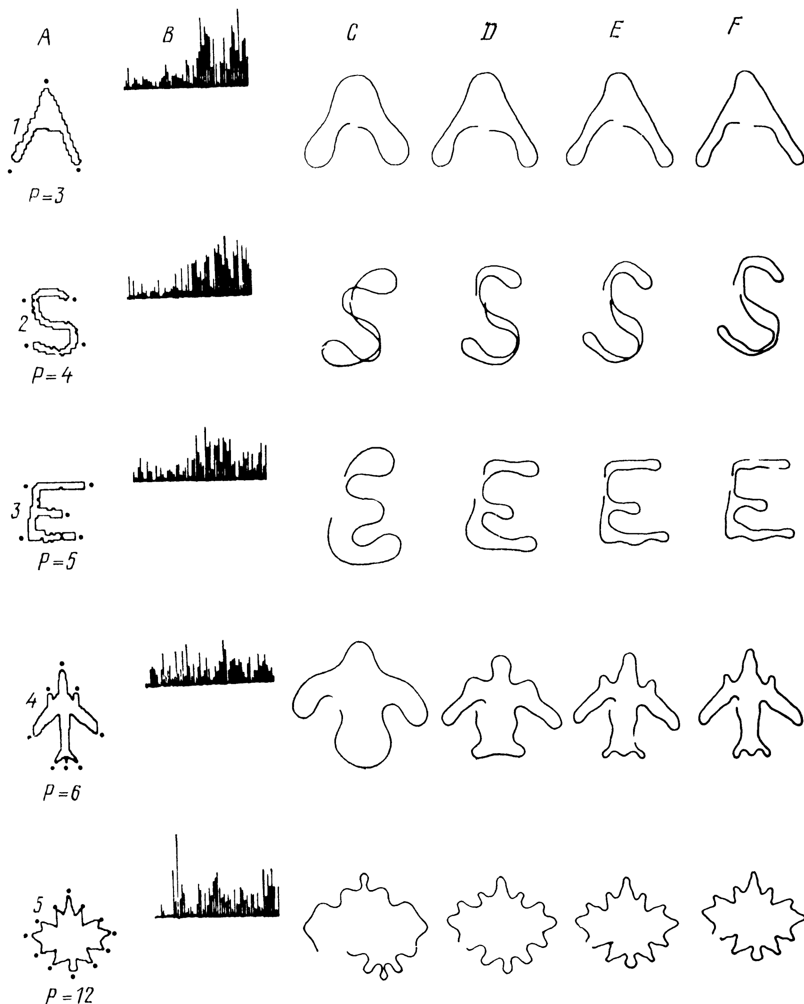


Рис. 18.4.4. Примеры разложений в ряд Фурье функции кривизны [35].

ностей смешанный момент (m, n)-го порядка совместной плотности вероятности $p(x, y)$ определяется как

$$M_{m,n} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^m y^n \, p(x, y) \, dx \, dy. \qquad (18.4.20)$$

Зная совместную характеристическую функцию

$$\Phi(\omega_x, \omega_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\{i(\omega_x x + \omega_y y)\} p(x, y)\, dx\, dy, \qquad (18.4.21)$$

смешанный момент можно определить как

$$M_{m, n} = i^{-(m+n)} \left. \frac{\partial^m \partial^n \Phi(\omega_x, \omega_y)}{\partial \omega_x^m \partial \omega_y^n} \right|_{\substack{\omega_x=0 \\ \omega_y=0}}. \qquad (18.4.22)$$

Смешанные центральные моменты можно получить, заменяя в выражении (18.4.20) x на $x - \eta_x$ и y на $y - \eta_y$, где η_x и η_y — маргинальные средние значения плотности величин x и y соответственно.

Эти классические соотношения теории вероятностей были применены Ху [36] и Алтом [37] для описания формы. Идея довольно проста. Совместная вероятность в выражениях (18.4.20) и (18.4.21) заменяется на функцию изображения $F(x, y)$. Форма объекта представляется несколькими моментами низкого порядка. Следует заметить, что выражение (18.4.21) непосредственно связано со стандартным непрерывным двумерным преобразованием Фурье, и поэтому следует ожидать тесной связи между моментами и компонентами спектра Фурье низкого порядка для функции изображения.

18.5. СЕГМЕНТАЦИЯ ПО ЯРКОСТИ

Сегментация изображения представляет собой разделение или разбиение изображения на области по сходству свойств их точек. Наиболее часто сегментацию проводят по яркости для одноцветного изображения и цветовым координатам для цветного изображения. В разд. 18.6 рассматривается сегментация, основанная на контурах; в разд. 18.7 описывается сегментация, при которой в качестве разделительного признака используется текстура, а разд. 18.8 посвящен сегментации по форме. Принятое здесь определение сегментации сильно ограничено: при сегментации не учитывается информация, характеризующая объекты в целом. Кроме того, в процессе сегментации производится лишь подразделение изображения, не делается попытки распознать отдельные сегменты, установить взаимные соотношения между ними и пометить их соответствующими символами. Этим важным этапам работы системы распознавания изображений посвящена гл. 20.

18.5.1. ПОРОГОВОЕ ОГРАНИЧЕНИЕ ПО ЯРКОСТИ

Многие изображения можно охарактеризовать тем, что они содержат некоторый интересующий нас объект достаточно однородной яркости на фоне другой яркости. Типичными примерами

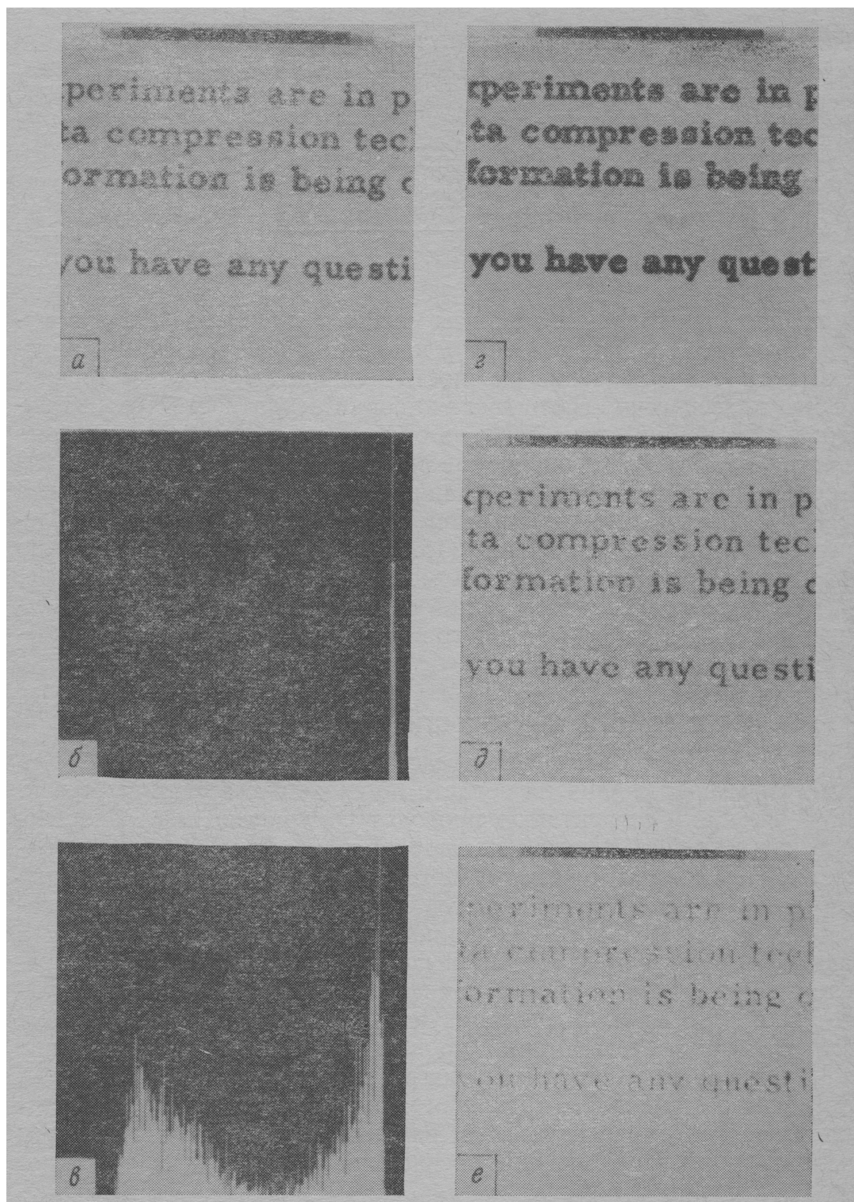


Рис. 18.5.1. Примеры сегментации машинописного текста методом порогового ограничения по яркости.

а — многоградационное изображение; *б* — гистограмма распределения яркости; *в* — гистограмма распределения яркости пограничных точек, выделенных оператором Лапласа; *г* — ограничение высоким порогом, $T = 170$; *д* — ограничение средним порогом, $T = 127$, *е* — ограничение низким порогом, $T = 25$

могут служить машинописные и рукописные тексты, медицинские пробы под микроскопом, самолеты на взлетной полосе. Для таких изображений яркость служит отличительным признаком, который можно использовать для локализации объекта. Если интересующий нас объект имеет белый цвет и расположен на черном фоне или наоборот, то определение точек объекта представляет собой тривиальную задачу установления порога по средней яркости. На практике, однако, встречаются определенные трудности, например когда наблюдаемое изображение подвержено воздействию шума и причем как на объекте, так и на фоне допускается некоторый разброс значений яркости. Другая часто встречающаяся трудность состоит в том, что фон может быть неоднородным.

На рис. 18.5.1, *а* приведена фотография представленного в цифровой форме машинописного текста, состоящего из темных букв на светлом фоне. На рис. 18.5.1, *б* показана гистограмма распределения яркости этого текста. Ожидаемая бимодальность гистограммы скрыта из-за относительно большого процента элементов фона. На рис. 18.5.1, *г*, *е* показан текст, выведенный на дисплей при разных порогах, причем всем точкам с яркостью выше порога соответствует уровень яркости, равный 255, а остальным точкам с яркостью ниже порога поставлен в соответствие нулевой уровень яркости. Эти фотографии иллюстрируют общую проблему, связанную с пороговым ограничением изображения. Если порог установить слишком низким, то отдельные части букв исчезнут (разорвана ножка буквы «р»). Наоборот, если порог установить слишком высоким, то в результате на объектах возникнут побочные детали (заполняется петелька буквы «е»). Наилучший компромисс обычно определяется экспериментальным путем. На рис. 18.5.2 приведен пример порогового ограничения по яркости многоградационного изображения, содержащего единственный значительный объект на нейтральном фоне.

Известно несколько аналитических подходов к пороговому ограничению по яркости. Один из методов состоит в установлении порога на таком уровне, при котором общая сумма элементов с подпороговой яркостью согласована с априорными вероятностями этих значений яркости [38]. Например, может быть известно, что черные знаки занимают 25% общей площади машинописной страницы. Тогда пороговый уровень для такого изображения необходимо установить таким, чтобы яркость четверти элементов была ниже порога. Другой подход к пороговому ограничению по яркости состоит в выборе порога, соответствующего минимуму бимодальной гистограммы, находящемуся между двумя ее пиками [39]. Определение этого минимального значения часто затруднено вследствие ступенчатости гистограммы. Поэтому участок гистограммы между пиками аппроксимируется некоторой аналитической функцией и находится ее минимум путем вычисления производных.

Например, пусть x и y — соответственно оси абсцисс и ординат на гистограмме. Тогда кривая второго порядка

$$y = ax^2 + bx + c, \tag{18.5.1}$$

где a, b и c — константы, обеспечивает простую аппроксимацию гистограммы в районе ее «долины». Минимум гистограммы наблюдается при $x = -b/2a$.

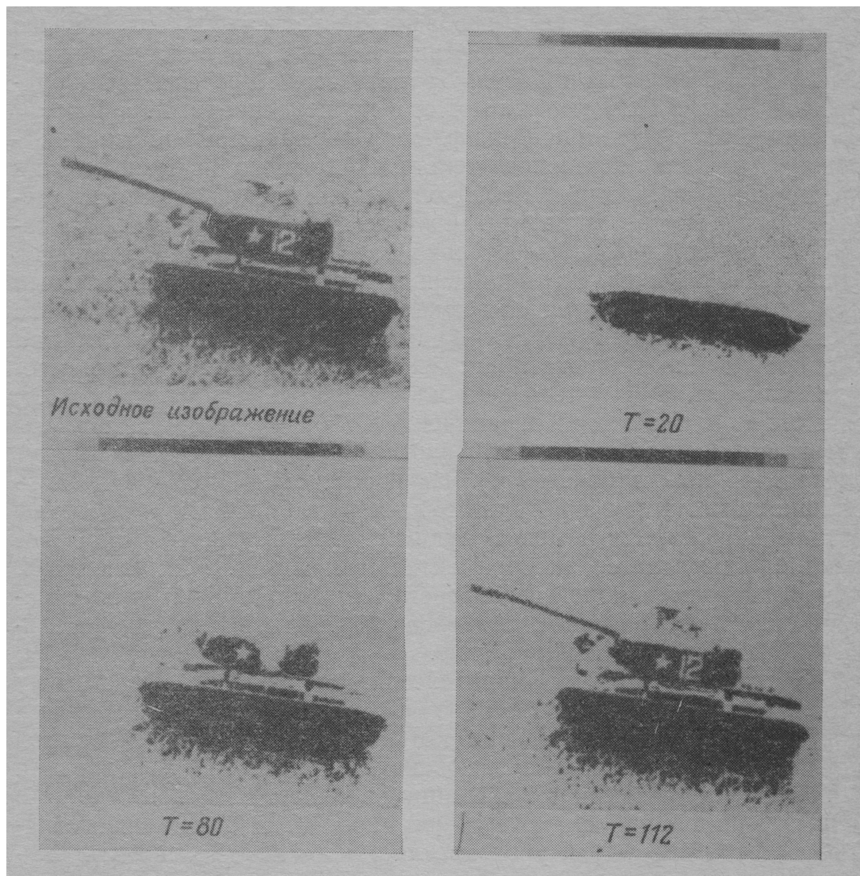


Рис. 18.5.2. Пример сегментации изображения «Танк» методом порогового ограничения по яркости.

Для определения порога яркости Уэшка, Нэгел и Розенфельд [40] предложили использовать оператор Лапласа. Для непрерывного изображения $F(x, y)$ оператор Лапласа

$$\nabla^2 F(x, y) = [\partial^2 F(x, y)/\partial x^2] + [\partial^2 F(x,y)/\partial y^2] \tag{18.5.2}$$

дает значения вторых частных производных этой функции по направлениям координатных осей. Рассмотрим область изображения в районе объекта, где яркость увеличивается с уровня низкого

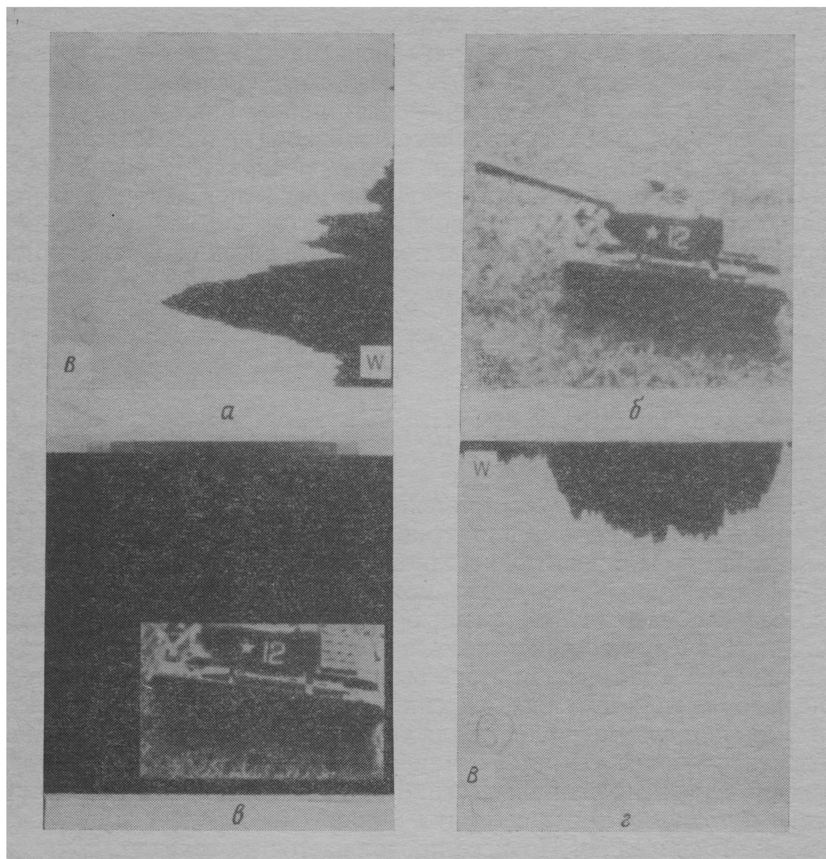


Рис. 18.5.3. Пример сегментации изображения «Танк» методом поперечных сечений профиля яркости.

а — гистограмма по строкам; *б* — исходное изображение, *в* — сегментация; *г* — гистограмма по столбцам.

«плато» до уровня более высокого «плато», соединенных наклонной поверхностью. На плоских участках лапласиан равен нулю, а вдоль наклонной поверхности — почти нулю. В области перехода от низкого «плато» лапласиан будет иметь большое положительное значение, а при переходе к высокому «плато» — большое отрицательное значение. Гистограмма, построенная с использованием лишь точек исходного изображения, которые соответствуют очень

высоким или очень низким значениям лапласиана, оказывается бимодальной с отчетливой «долиной» между пиками. Определению порога яркости могут помочь и другие процедуры выделения перепадов, которые были рассмотрены в гл. 17.

Если фон на изображении неоднороден, то необходимо сделать так, чтобы порог по яркости адаптировался к ее среднему уровню [41, 42]. Это можно выполнить, разбивая изображения на маленькие фрагменты и устанавливая свои пороги для каждого фрагмента. Пороги для каждого элемента изображения можно затем определить путем интерполяции между центрами фрагментов.

Области изображения иногда можно выделить путем усреднения поперечных сечений профиля яркости изображения [43, 44]. Горизонтальные и вертикальные поперечные сечения определяются следующим образом:

$$H(k) = (1/N) \sum_{j=0}^{N-1} F(j, k), \tag{18.5.3}$$

$$V(j) = (1/N) \sum_{k=0}^{N-1} F(j, k). \tag{18 5.4}$$

Рис. 18.5.3 иллюстрирует применение метода сегментации по поперечным сечениям профиля яркости для определения размеров объекта на поле с достаточно однородным фоном.

Эффективную сегментацию можно выполнить для некоторых классов изображений с помощью рекуррентного метода порогового ограничения по яркости, который предложили Томита, Ясида и Цудзи [45]. На первом этапе выполняется пороговое ограничение изображения путем нахождения минимального значения между модами гистограммы яркости для того, чтобы отделить более яркие области от более темных. Затем формируются гистограммы для каждой сегментированной части. Если эти гистограммы неунимодальны, то полученные сегменты можно снова подвергнуть пороговому ограничению. Процесс продолжается до тех пор, пока гистограммы для всех сегментов не станут унимодальными или же не будет достигнут некоторый желаемый уровень сегментации.

18.5.2. МНОГОМЕРНОЕ ПОРОГОВОЕ ОГРАНИЧЕНИЕ

Идею порогового ограничения по яркости можно распространить на сегментацию цветных и спектрозональных изображений. Олендер [46] разработал схему сегментации для естественных цветных изображений, основанную на многомерном пороговом ограничении цветных изображений, представленных в системах координат цвета *RGB*, *YIQ*, а также в системе нестандартных цветовых координат, получивших неудачные названия интенсив-

ность, цветовой тон и насыщенность [1]). На рис. 18.5.4 дан пример гистограмм этих девяти цветовых координат для сцены, представляющей собой комнату с обстановкой для отдыха. Эти гистограммы были получены для тех частей исходной сцены, на которых отсутствует текстура. Важный предварительный этап сегментации текстурной части сцены, которая рассматривается в разд. 18.7, необходим для того, чтобы исключить ложную сегментацию областей с однородной текстурой на множество изолированных частей. Если не все гистограммы признаков унимодальны, то применяется рассчитанная специально на этот случай процедура определения наилучшего признака и наилучшего уровня, по которому выполняется пороговое ограничение этого признака. Первым кандидатом является интенсивность [2]), позволяющая разделить изображения на черные и белые области. Следующий кандидат выбирается на основании приоритета, зависящего от уровня контраста и положения мод на гистограмме. После того как пороговый уровень определен, изображение разбивается на части. Затем процедура повторяется для каждой части до тех пор, пока получающиеся гистограммы признаков не станут унимодальными или сегментация не достигнет приемлемого для наблюдателя уровня разделения. Модельный эксперимент показал, что метод сегментации Олендера с использованием многомерного порогового ограничения в сочетании с различением текстур оказывается весьма эффективным. Однако большая часть управляющих операций при моделировании выполнялась человеком-оператором; для настройки параметров процедуры требуется человеческое суждение, основанное на визуальном анализе результатов пробных экспериментов с разными пороговыми значениями.

---

[1]) Эти не описанные в литературе цветовые координаты, введенные Тененбаумом (Стэнфордский исследовательский институт) и широко используемые при анализе сцен, определяются следующим образом:
Интенсивность

$$I = R + G + B.$$

Насыщенность

$$S = 1 - 3 \min (r, g, b),$$

где $r = R/I$, $g = G/I$, $b = B/I$.
Цветовой тон

$$\Phi = \arccos \left[\frac{2r - g - b}{\sqrt{6} \left[\left(r - \frac{1}{3}\right)^2 + \left(g - \frac{1}{3}\right)^2 + \left(b - \frac{1}{3}\right)^2 \right]^{1/2}} \right],$$

где $\theta = \Phi$ по модулю 180°.

[2]) По-видимому, сигнал яркости Y больше подходит, поскольку яркость — психофизическая характеристика, тогда как интенсивность, определяемая в указанной работе как сумма $R + G + B$, таковой не является.

В методе сегментации Олендера предполагается, что девять признаков взаимосвязаны. Сигналы Y, I, Q представляют собой линейную комбинацию сигналов R, G, B, а сигналы интенсивности,

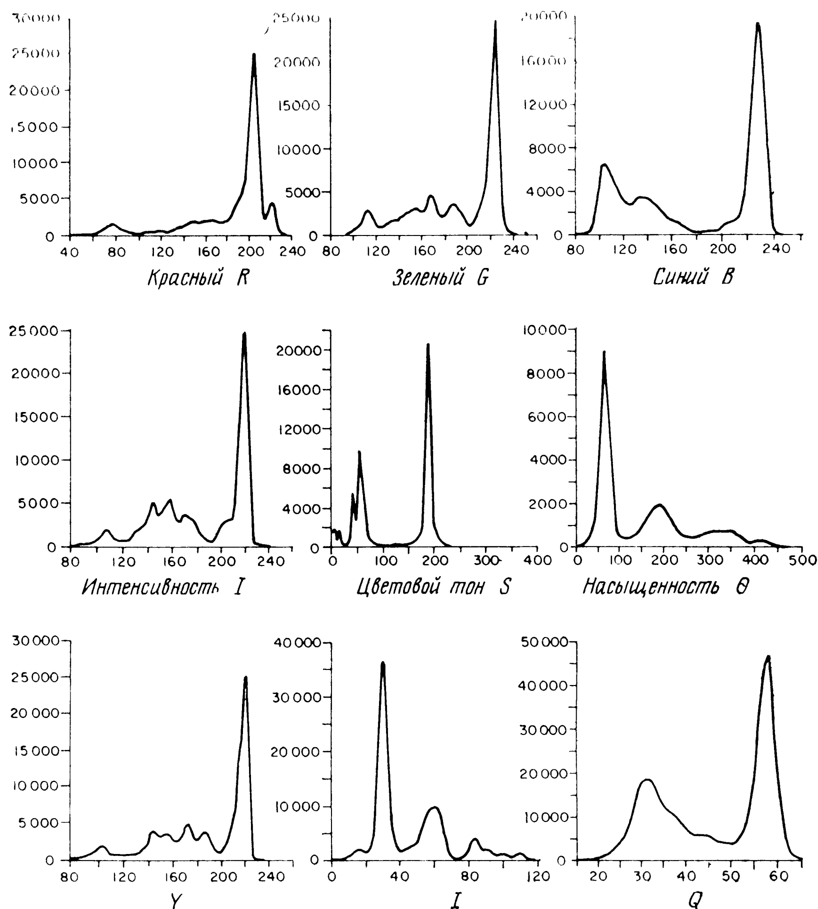


Рис. 18.5.4. Типичные гистограммы признаков для сегментации цветных изображений [46].

цветового тона и насыщенности образуют нелинейные комбинации R, G, B. Это замечание порождает несколько вопросов. Каковы наилучшие для сегментации типы линейных или нелинейных преобразований? Сколько значений признаков следует использовать? Что является наилучшей формой порогового ограничения? Возможно, ответы на эти вопросы появятся в процессе изучения мето-

дов кластерного анализа, используемого в задачах распознавания образов [11, 47—49].

На самом деле одномерные гистограммы признаков являются маргинальными гистограммами некоторой совместной многомерной гистограммы. Методы кластерного анализа позволяют указать границы областей принятия решений в многомерном пространстве и использовать их при сегментации. Такой подход дает возможность наиболее полно использовать все признаки с учетом их взаимной корреляции. Можно также применить метод линейного дискриминантного анализа Фишера [11, стр. 114], чтобы одновременно спроектировать все значения признаков на прямую и воспользоваться группировкой спроектированных данных для выполнения одномерного порогового ограничения.

18.5.3. НАРАЩИВАНИЕ ОБЛАСТЕЙ

Наращивание областей — один из наиболее простых для понимания подходов к сегментации по яркости: соседние элементы с одинаковыми яркостями группируются вместе и образуют область. Однако на практике для достижения приемлемых результатов на процесс наращивания областей требуется наложить ограничения, причем некоторые из них достаточно сложные.

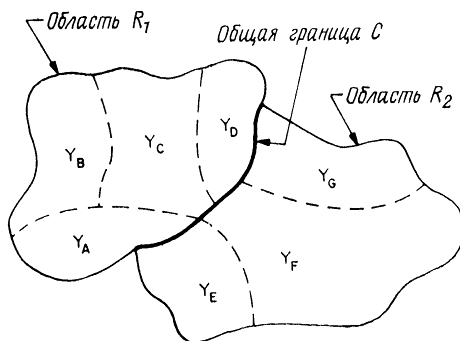


Рис. 18.5.5. Наращивание областей.

Брайс и Феннема [50] разработали и испытали метод наращивания областей, основанный на выборе простых правил роста. На первом этапе процесса пары квантованных элементов изображения объединяются в группы, называемые «атомами», если они обладают одинаковой яркостью и являются четырехсвязными. Затем по двум эвристическим правилам удаляются слабые границы между атомами. Обратимся к рис. 18.5.5. Пусть R_1 и R_2 есть две смежные области, полученные на предыдущем шаге слияния с периметрами P_1 и P_2. Заметим, что после выполнения начальных этапов наращивания может оказаться, что некоторая область содержит ранее слитые подобласти с различными значениями яркости. Пусть C означает длину общей границы, а D — длина той части границы, для которой разность яркостей ΔY граничащих элементов меньше

некоторого коэффициента значимости ε_1. Тогда области R_1 и R_2 сливаются, если

$$D/\min\{P_1, P_2\} > \varepsilon_2, \quad (18.5.5)$$

где ε_2 — константа, обычно равная $^1/_2$. Это эвристическое правило предотвращает слияние смежных областей приблизительно одинакового размера, но допускает поглощение меньших областей бóльшими. По второму правилу происходит ликвидация общих слабых границ, остающихся после применения первого правила. Смежные области сливаются, если

$$D/C > \varepsilon_3, \quad (18.5.6)$$

где ε_3 — константа, приблизительно равная $^3/_4$. Применение только второго правила ведет к чрезмерному слиянию областей.

Наращивание областей по методу Брайса и Феннема обеспечивает достаточно точную сегментацию простых сцен с небольшим числом объектов и без текстуры [50, 51], однако для более сложных сцен этот метод не дает хороших результатов. В работе [52] сделана попытка усовершенствовать идею наращивания областей, введя ограничения на слияние, основанные на бейесовской оценке признаков для каждой области. Тененбаум и Вейл [53] использовали семантические знания для управления процессом слияния. Оба метода дают лучшие результаты, чем метод Брайса и Феннема, однако они очень сложны.

18.6. КОНТУРНАЯ СЕГМЕНТАЦИЯ

Обычно анализ изображений включает такие операции, как получение внешнего контура изображенных объектов и запись координат точек этого контура. Часто требуется получить внешний контур в виде замкнутой кривой или совокупности отрезков дуг. Имеется три общих подхода к представлению границ объекта: аппроксимация кривых, прослеживание контуров и связывание точек перепадов.

Аппроксимация кривых

Методы аппроксимации кривых, описанные в разд. 18.3, можно использовать для контурной сегментации путем подбора аналитически заданной кривой к совокупности точек контурного препарата, полученного с помощью детектора перепадов. Если имеется априорная информация относительно ожидаемой формы объекта, например известно, что это прямоугольник или окружность, то в качестве аппроксимирующей кривой нужно взять прямоугольник или окружность. Или же можно попытаться применить аппроксимацию полиномами, используя описанные в разд. 18.3 итеративные методы подбора прямых линий. Метод подбора кривых доста-

точно хорошо подходит для объектов с простой структурой. Трудности возникают, когда изображение содержит множество перекрывающихся объектов и точки перепадов образуют ветвистые структуры.

Прослеживание контуров

Прослеживание контуров напоминает поведение жука, обходящего препятствие. В случае двоичного изображения, например

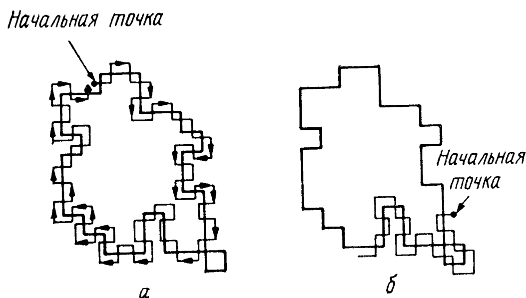


Рис. 18.6.1. Примеры прослеживания внешних контуров.
а — начальная точка расположена вверху слева; *б* — начальная точка расположена внизу справа.

показанного на рис. 18.6.1, *а*, воображаемый жук начинает свой путь на белом поле и движется по направлению к области черных элементов изображения. После того как жук пересечет черный элемент, он поворачивается налево и переходит к следующему элементу. Если этот элемент черный, жук снова поворачивается налево, если же элемент оказывается белым, то жук поворачивается направо. Эта процедура продолжается до тех пор, пока жук не вернется в исходную точку. Декартовы координаты точек перехода с черного на белое или с белого на черное дают местоположение границы. На рис. 18.6.1, *а* выступающий элемент в нижнем правом углу объекта, обладающий восьмисвязностью со своим ближайшим соседом, не включен в границу. Заметим, однако, что этот выступающий элемент на рис. 18.6.1, *б* включен в границу объекта, но здесь начальная точка передвинута. Таким образом, определение границы зависит от точки начала движения. Другая трудность связана с наличием «дыр» внутри объекта, которые жук может пропустить. Проблемы такого характера можно преодолеть, снабдив жука некоторой памятью и интеллектом, что позволит ему запоминать свои прошлые шаги и возобновлять прослеживание, если данный путь оказывается ошибочным. Описанный алгоритм чаще всего применяется к яркостным изображениям, под-

вергнутым пороговому ограничению, но данную идею легко можно распространить и на многоградационные изображения. Один из подходов состоит в том, что граничная точка отмечается, если разница яркостей соседних элементов изображения достаточно велика. В результате жук формирует контуры по ходу своего движения.

Соединение точек перепадов

Объект или область изображения, которую желательно выделить, обычно отделяется от соседних частей изображения перепадом яркости. Поэтому соединение точек перепадов можно использовать для сегментации.

Метод соединения точек перепадов, разработанный Робертсом [54], основан на тех же основных принципах, что и большинство

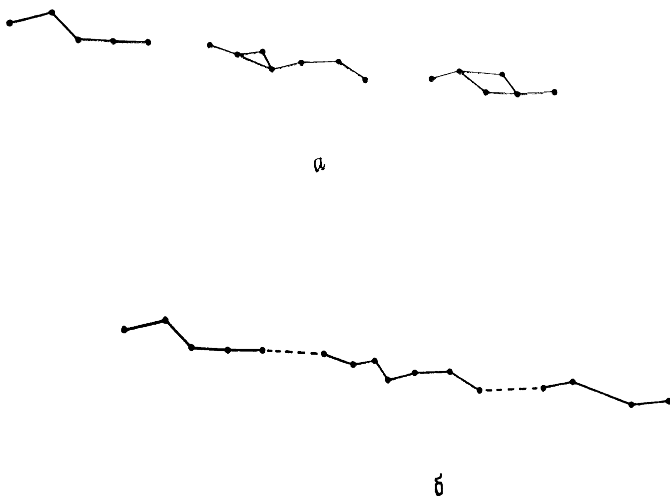


Рис. 18.6.2. Примеры соединения точек перепадов по правилу Робертса.
а — соединение точек перепадов; *б* — исключение множественных соединений и замена их мостиками из прямых линий.

других методов такого связывания. В методе Робертса аналоговые значения градиентов (разностей яркости соседних элементов) анализируются блоками размером 4×4 элемента. Элемент с наибольшим в блоке значением модуля градиента считается пробной контурной точкой, если это значение больше порога. Затем к значениям градиента около этой пробной точки подбираются линии длиной в пять элементов с ориентацией «север», «восток», «юг» и «запад». Если отношение наилучшей аппроксимации к наихудшей больше второго порога, то пробная контурная точка объявляется

действительной и ей приписывается ориентация, соответствующая наилучшей аппроксимации. Далее к парам контурных точек подбираются прямые линии, если эти точки расположены в смежных блоках размером 4×4 элемента и если направление каждой линии находится в диапазоне ±23° относительно ориентации контура в каждой контурной точке. Точки, не удовлетворяющие критериям соединения, отбрасываются. Типичная граница, полученная на этом этапе, будет содержать, как видно из рис. 18.6.2, *а*, разрывы, и множественные соединения точек. Маленькие треугольники исключаются вычерчиванием наибольшей стороны, а маленькие четырехугольники заменяются их наибольшей диагональю (рис. 18.6.2, *б*). Короткие выступающие линии также уничтожаются. На этом этапе короткие разрывы заменяются мостиками из прямых линий. Этот вид соединения точек перепада можно ис-

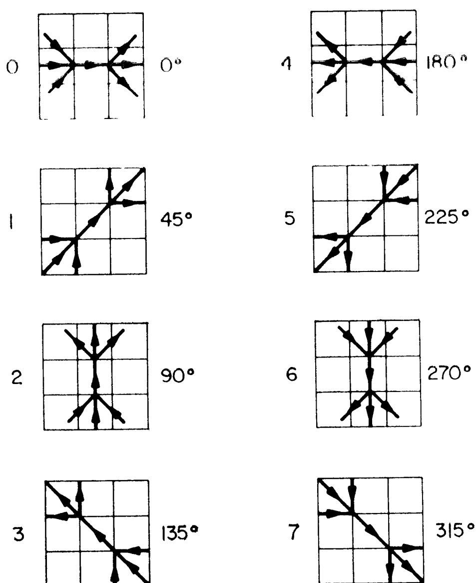


Рис. 18.6.3 Правила соединения точек перепадов [56].

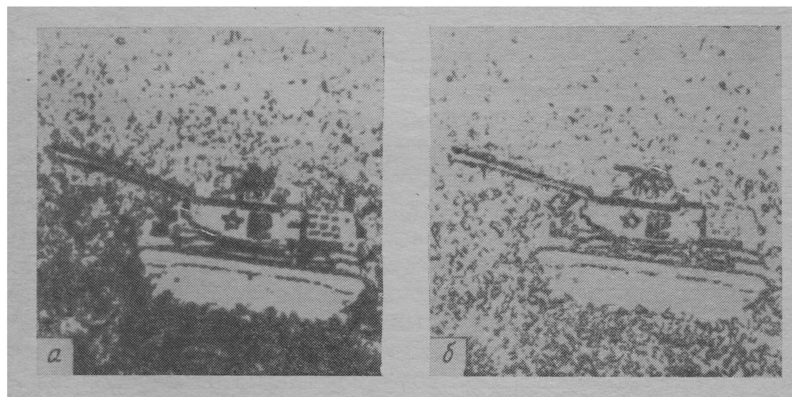


Рис. 18.6.4 Пример построения контуров с использованием алгоритма соединения точек перепада [56].
а — без соединения; *б* — с соединением.

пользовать для широкого класса детекторов перепадов. В работе [55] подобный метод применен для соединения точек, полученных с помощью детектора перепадов Хюккеля.

Робинсон [56] предложила простой, но эффективный алгоритм соединения, при использовании которого точки перепада, полученные с помощью детектора, дающего возможность определить ориентацию перепада в системе восьми курсовых направлений, исследуются по блокам из 3×3 элементов (рис. 18.6.3). Точка перепада в центре области считается действительной контурной точкой, если ее соседи по соответствующим направлениям принадлежат к перепадам надлежащей ориентации. Рис. 18.6.4 иллюстрирует возможность данного способа удалять явно посторонние перепады фона. Применение окон большего размера было бы полезно, однако с увеличением размера окна количество возможных связей действительного контура будет быстро расти.

18.7. ТЕКСТУРНАЯ СЕГМЕНТАЦИЯ

Уже давно стало ясно, что текстура была бы ценным признаком для сегментации изображений. Однако применению этой идеи на практике препятствовало до сих пор отсутствие надежных и эффективных способов обнаружения и измерения параметров текстуры.

Один из подходов к текстурной сегментации, разработанный Розенфельдом и др. [57—59], состоит в подсчете некоторой меры зернистости текстуры во всех точках изображения с последующим обнаружением изменений этой меры. Фактически исходное изображение подвергается предварительной обработке с тем, чтобы преобразовать его в яркостное изображение. Основная трудность, связанная с этим подходом, состоит в том, что параметры текстуры измеряются в некотором окне и, следовательно, их значения, полученные в окрестности границы между текстурными областями, являются усредненными. В результате трудно точно локализовать границу между текстурными областями.

Другой подход к текстурной сегментации заключается в обнаружении переходной зоны между областями с различной текстурой. Основная идея обнаружения изменения текстуры идентична идее обнаружения яркостного перепада: различие между текстурными областями усиливается во всех точках изображения, а затем контрастированный препарат подвергается пороговому ограничению для локализации текстурных перепадов. Томпсон [60] предложил средства контрастирования текстуры, аналогичные градиенту Робертса, представленному в разд. 17.4. Параметры текстуры вычисляются для каждого из четырех сканируемых по изображению смежных блоков размером $W \times W$ элементов и находится сумма модулей перекрестных разностей, а затем выполняется по-

роговое ограничение для локализации значительных изменений текстуры. Этот метод можно обобщить так, чтобы выполнять вычисления по смежным окнам, объединенным в группы 3×3. Тогда значения признаков текстуры, вычисленные в каждом окне, можно объединить некоторым линейным или нелинейным способом, аналогичным методам контрастирования перепадов яркости по области 3×3, описанным в разд. 17.4.

Цуккер, Розенфельд и Девис [61] предложили гистограммный пороговый метод текстурной сегментации, основанный на разработанном Цудзи и Томита [62] способе анализа текстуры. При использовании этого метода значение признака текстуры подсчитывается в каждой точке изображения по градиенту текстуры с использованием алгоритма подавления доминирующим соседом, который был описан в разд. 17.4. Затем по полученным модифицированным данным градиента формируется гистограмма. Если гистограмма многомодальна, то пороговое ограничение градиента по минимуму гистограммы, находящемуся между ее модами, должно приводить к сегментации текстурных областей. Этот процесс можно повторять для отдельных частей изображения до тех пор, пока не будет завершена сегментация.

18.8. СЕГМЕНТАЦИЯ ФОРМЫ

В задачах распознавания изображений часто оказывается полезным разбить объект сложной формы на связанный набор частей простой, легко описываемой формы. Например, букву «Т» можно

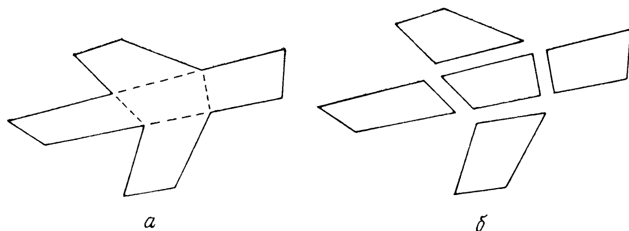


Рис. 18.8.1. Пример сегментации формы.
а — исходный объект, *б* — сегментированный объект.

разбить на связанные прямоугольники. Сегментацию формы можно выполнить в два этапа. Сначала объект произвольной формы аппроксимируется набором связанных прямолинейных или криволинейных отрезков. Затем аппроксимированная форма подвергается сегментации по точкам перегиба. Рис. 18.8.1 иллюстрирует сегментацию объектов, имеющих форму многоугольников. Основные правила сегментации достаточно просты: ближайшие точки

вогнутости соединяют, образуя «перешейки», которые затем удаляют из основного тела; процесс продолжается до тех пор, пока все части не окажутся выпуклыми. Фэн и Павлидис [63] разработали алгоритмы эффективной сегментации фигур, имеющих форму многоугольника.

ЛИТЕРАТУРА

1. Rosenfeld A., Connectivity in Digital Pictures, *JACM*, **17**, 1, 146—160 (January 1970).
2. Rosenfeld A., Picture Processing by Computer, Academic Press, New York, 1969. [Имеется перевод: Розенфельд А., Распознавание и обработка изображений. — М.: Мир, 1972].
3. Golay M. J. E., Hexagonal Pattern Transformation, *IEEE Trans. Computers*, **C-18**, 8, 733—740 (August 1969).
4. Preston K., Jr., Feature Extraction by Golay Hexagonal Pattern Transforms, *IEEE Trans. Computers*, **C-20**, 9, 1007—1014 (1971).
5. Preston K., Jr., The CELLSCAN System, a Leucocyte Pattern Analyzer, Western Joint Computer Conference, May 1961, pp. 173—183.
6. Unger S. H., A Computer Oriented Towards Spatial Problems, *Proc. IRE*, **46**, 10, 1744—1750 (October 1958).
7. Levialdi S., CLOPAN: A Closed Pattern Analyzer, *Proc. IRE*, **115**, 6, 879—880 (June 1968).
8. Kruse B., A Parallel Picture Processing Machine, *IEEE Trans. Computers*, **C-22**, 12, 1075—1087 (December 1973).
9. Duff M. J. B. et al., A Cellular Logic Array for Image Processing, *Pattern Recognition*, **5**, 229—247 (1973).
10. Blum H., A Transformation for Extracting New Descriptors of Shape, in: Symposium Models for Perception of Speech and Visual Form, Weiant Whaten-Dunn, Ed., MIT Press, Cambridge, Mass., 1967.
11. Duda R. O., Hart P. E., Pattern Classification and Scene Analysis, Wiley-Interscience, New York, 1973. [Имеется перевод: Дуда Р., Харт П., Распознавание образов и анализ сцен. — М.: Мир, 1976.]
12. Calabi L., Hartnett W. E., Shape Recognition, Prairie Fires, Convex Deficiencies and Skeletons, *Am. Math. Monthly*, **75**, 4, 335—342 (April 1968).
13. Mott-Smith J. C., Medial Axis Transforms, in: Picture Processing and Psychopictorics, Lipkin B. S., Rosenfeld A., Eds., Academic Press, New York, 1970.
14. Montanari U., A Method for Obtaining Skeletons Using a Quasi-Euclidean Distance, *JACM*, **15**, 600—624 (October 1968).
15. Montanari U., Continuous Skeletons from Digitized Pictures, *JACM*, **16**, 4, 534—549 (October 1969).
16. McCalla T. R., Introduction to Numerical Methods and FORTRAN Programming, Wiley, New York, 1967.
17. Ramer U., An Iterative Procedure for the Polygonal Approximation of Plane Curves, *Computer Graphics and Image Processing*, **1**, 3, 244—256 (November 1972).
18. Pavlidis T., Horowitz S. L., Segmentation of Plane Curves, *IEEE Trans. Computers*, **C-23**, 8, 860—870 (August 1974).
19. Hough P. V. C., Method and Means for Recognizing Complex Patterns, U. S. Patent 3069654, December 18, 1962.
20. Duda R. O., Hart P. E., Use of the Hough Transformation to Detect Lines and Curves in Pictures, *Commun. ACM*, **15**, 1, 11—15 (January 1972).
21. O'Gorman F., Clowes M. B., Finding Picture Edges Through Collinearity of Feature Points, *IEEE Trans. Computers*, **C-25**, 4, 449—456 (April 1976).

22. Kolers P. A., The Role of Shape and Geometry in Picture Recognition, in: Picture Processing and Psychopictorics, Lipkin B. S., Rosenfeld A., Eds., Academic Press, New York, 1970.
23. Attneave F., Physical Determinants of the Judged Complexity of Shape, *J. Experimental Psychology*, **53**, 221—227 (1957).
24. Attneave F., Arnoult M., The Quantitative Study of Shape and Pattern Perception, *Psychology Bull.*, **53**, 452—471 (1956).
25. Rosenfeld A., Pflatz J. L., Distance Functions on Digital Pictures, *Pattern Recognition*, **1**, 33—62 (July 1968).
26. Greanis E. C., et al., The Recognition of Handwritten Numerals by Contour Analysis, *IBM J. Res. Devel.*, **7**, 1, 14—21 (January 1963).
27. Fischler M. A., Machine Perception and Description of Pictorial Data, Proceedings International Joint Conference on Artificial Intelligence, Walker D. E., Norton L. M., Eds., May 1969, pp. 629—639.
28. Gray S. B., Local Properties of Binary Images in Two Dimensions, *IEEE Trans. Computers*, **C-20**, 5, 551—561 (May 1971).
29. Sklansky J., Recognizing Convex Blobs, Proceedings International Joint Conference on Artificial Intelligence, Walker D. E., Norton L. M., Eds., May 1969, pp. 107—116.
30. Sklansky J., Cordella L. P., Levialdi S., Parallel Detection of Concavities in Cellular Blobs, *IEEE Trans. Computers*, **C-25**, 2, 187—196 (February 1976).
31. Duda R. O., Image Segmentation and Description (неопубликованные данные), 1975.
32. Cosgriff R. L., Identification of Shape, Ohio State University Research Foundation, Columbus, Ohio, Report 820-11, ASTIA AD 254 792, December 1960.
33. Brill E. L., Character Recognition Via Fourier Descriptors, WESCON Convention Record, Paper 25/3, Los Angeles, 1968.
34. Zahn C. T., Roskies R. Z., Fourier Descriptors for Plane Closed Curves, *IEEE Trans. Computers*, **C-21**, 3, 269—281 (March 1972).
35. Bennett J. R., MacDonald J. S., On the Measurement of Curvature in a Quantized Environment, *IEEE Trans. Computers*, **C-25**, 8, 803—820 (August 1975).
36. Hu M. K., Visual Pattern Recognition by Moment Invariants, *IRE Trans. Inf. Theory*, **IT-8**, 2, 179—187 (February 1962).
37. Alt F. L., Digital Pattern Recognition by Moments, *JACM*, **9**, 2, 240—258 (April 1962).
38. Doyle W., Operations Useful for Similarity-Invariant Pattern Recognition, *JACM*, **9**, 2, 259—267 (April 1962).
39. Prewitt J. M. S., Mendelsohn M. L., The Analysis of Cell Images, *Ann. N. Y. Acad. Sci.*, **128**, 1035—1053 (1966).
40. Weszka J. S., Nagel R. N., Rosenfeld A., A Threshold Selection Technique, *IEEE Trans. Computers*, **C-23**, 12, 1322—1326 (December 1974).
41. Bartz M. R., The IBM 1975 Optical Page Reader, Part II: Video Thresholding System, *IBM J. Res. Devel.*, **12**, 354—363 (September 1968).
42. Chow C. K., Kaneko T., Boundary Detection of Radiographic Images by a Threshold Method, in: Frontiers of Pattern Recognition, Watanabe S., Ed., Academic Press, New York, 1972.
43. Becker H. C. et. al., Digital Computer Determination of a Medical Diagnostic Index Directly from Chest X-ray Images, *IEEE Trans. Biomed. Eng.*, **BME-11**, 3, 67—72 (July 1964).
44. Kruger R. P. et al., Radiographic Diagnosis via Feature Extraction and Classification of Cardiac Size and Shape Descriptors, *IEEE Trans. Biomed. Eng.*, **BME-19**, 3, 174—186 (May 1972).
45. Tomita F., Yachida M., Tsuji S., Detection of Homogeneous Regions by Structural Analysis, Proceedings International Joint Conference on Artificial Intelligence, Stanford, California, August 1973, pp. 564—571.
46. Ohlander R. B., Analysis of Natural Scenes, Carnegie-Mellon University, Department of Computer Science, Ph. D. Thesis, April 1975.

47. Andrews H. C., Introduction to Mathematical Techniques in Pattern Recognition, Wiley-Interscience, New York, 1972.
48. Meisel W. S., Computer-Oriented Approaches to Pattern Recognition, Academic Press, New York, 1972.
49. Haralick R. M., Kelly G. L., Pattern Recognition with Measurement Space and Spatial Clustering for Multiple Images, *Proc. IEEE*, **57**, 4, 654—665 (April 1969). [Имеется перевод: Хaралик Р. М., Келли Г. Л. Использование методов распознавания образов и автоматической классификации для анализа многокомпонентных изображений. — ТИИЭР, 1969, т. 57, № 4, с. 316—339.]
50. Brice C. R., Fenema C. L., Scene Analysis Using Regions, *Artificial Intelligence*, **1**, 205—226 (1970). [Имеется перевод: Брайс К. Р., Феннема К. Л. Анализ сцены при помощи выделения областей. — В сб. Интегральные роботы. — М.: Мир, 1975, т. 2, с. 136—159.]
51. Barrow H. G., Popplestone R. J., Relational Descriptions in Picture Processing, in: Machine Intelligence, Vol. 6, Meltzer B., Michie D., Eds., University Press, Edinburgh, 1971, pp. 377—396.
52. Yakimovsky Y., Scene Analysis Using a Semantic Base for Region Growing, Report AIM-209, Stanford University, 1973.
53. Tenenbaum J. M., Weyl S., A Region Analysis Subsystem for Interactive Scene Analysis, Proceedings Fourth International Joint Conference on Artificial Intelligence, September 1975. [Имеется перевод: Тененбаум Д., Вейль С. Подсистема анализа областей для интерактивного анализа сцен. Труды IV Международной объединенной конференции по искусственному интеллекту. — Тбилиси, 1975, т. 8, с. 8.153—8.165.]
54. Roberts L. G., Machine Perception of Three Dimensional Solids, in: Optical and Electro-Optical Information Processing, Tippett J. T., et al., Eds., MIT Press, Cambridge, Mass., 1965.
55. Nevatia R., Locating Object Boundaries in Textured Environments, *IEEE Trans. Computers*, **C-25**, 11, 1170—1175 (November 1976).
56. Robinson G. S., Detection and Coding of Edges Using Directional Masks, Proceedings SPIE Conference on Advances in Image Transmission Techniques, San Diego, August 1976.
57. Rosenfeld A., Thurston M., Edge and Curve Detection for Visual Scene Analysis, *IEEE Trans. Computers*, **C-20**, 5, 562—569 (May 1971).
58. Rosenfeld A., Thurston M., Lee Y. H., Edge and Curve Detection: Further Experiments, *IEEE Trans. Computers*, **C-21**, 7, 677—715 (July 1972).
59. Hayes K. C., Jr., Shah A. N., Rosenfeld A., Texture Coarseness: Further Experiments, *IEEE Trans. Systems, Man, and Cybernetics* (Correspondence), **SMC-4**, 5, 467—472 (September 1974).
60. Thompson W. B., Textural Boundary Analysis, University of Southern California, Image Processing Institute, Report USCIPI 620, 124—134 (September 1975).
61. Zucker S. W., Rosenfeld A., Davis L. S., Picture Segmentation by Texture Discrimination, *IEEE Trans. Computers*, **C-24**, 12, 1228—1233 (December 1975).
62. Tsuji S., Tomita F., A Structural Analyzer for a Class of Textures, *Computer Graphics and Image Processing*, **2**, 3/4, 216—231 (December 1973).
63. Feng H.-Y. F., Pavlidis T., Decomposition of Polygons into Simpler Components: Feature Generation for Syntactic Pattern Recognition, *IEEE Trans. Computers*, **C-24**, 6, 636—650 (June 1975).

Глава 19

ОБНАРУЖЕНИЕ ОБЪЕКТОВ И СОВМЕЩЕНИЕ (ПРИВЯЗКА) ИЗОБРАЖЕНИЙ

Настоящая глава охватывает две относящиеся к анализу изображений задачи: обнаружение объектов и совмещение изображений. Обнаружение изображений связано с установлением наличия объектов, относительно которых предполагается, что они имеются на картинке. Совмещение изображений состоит в нахождении поточечного соответствия (привязке) двух видов одной сцены.

19.1. СОПОСТАВЛЕНИЕ С ЭТАЛОНОМ

Один из основных способов обнаружения объектов на изображении состоит в сопоставлении с эталоном. При этом эталон интересующего нас объекта сравнивается со всеми неизвестными объектами, находящимися на изображении [1—4]. Если сходство между неизвестным объектом и эталоном достаточно велико, то этот объект помечается как соответствующий эталонному объекту.

В качестве простого примера сопоставления с эталоном рассмотрим набор двоичных фигур, образованных черными линиями на белом фоне (рис. 19.1.1, *а*). В этом примере задача состоит в том, чтобы обнаружить на изображении равнобедренный прямоугольный треугольник и определить его местоположение. На рис. 19.1.1, *б* показан эталонный треугольник. Толщина сторон выбрана так, чтобы обеспечить инвариантность результатов по отношению к небольшим изменениям формы искомого треугольника и при этом не слишком много потерять в точности его локализации. При работе эталон последовательно перемещается по полю изображения и исследуется его сходство с различными участками изображения.

Полное совпадение эталона с какой-либо частью изображения бывает редко из-за действия шумов и искажений, вызванных пространственной дискретизацией и квантованием яркости, а также вследствие отсутствия априорной информации относительно точной формы и структуры объекта, который требуется обнаружить. Поэтому обычно с помощью некоторой конкретной меры различия $D(m, n)$ между эталоном и изображением в точке (m, n) указывают на наличие выделенного объекта там, где это различие меньше

некоторого установленного порога $L_D(m, n)$. Как правило, порог выбирается постоянным для всех точек изображения. Обычно в качестве меры различия берется среднеквадратическая ошибка, определяемая как

$$D(m, n) = \sum_j \sum_k [F(j, k) - T(j-m, k-n)]^2, \qquad (19.1.1)$$

где $F(j, k)$ — элемент массива изображения, на котором производится поиск, а $T(j, k)$ — элемент эталонного массива. Поиск,

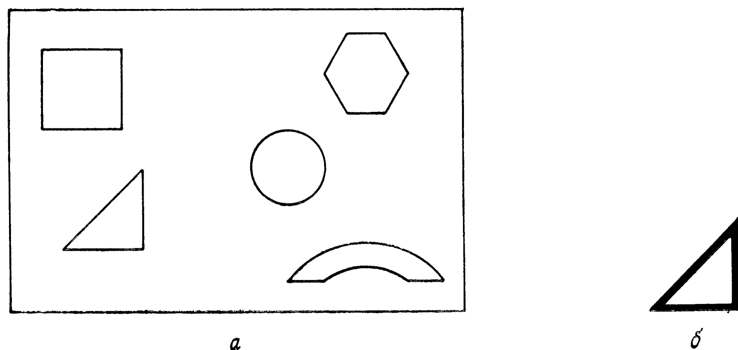


Рис. 19.1.1. Сопоставление с эталоном.
а — множество объектов; *б* — эталон для треугольника.

конечно, ограничен областью перекрытия смещенного эталона и изображения. Считается, что имеется сходство с эталоном в точке с координатами (m, n), если

$$D(m, n) < L_D(m, n). \qquad (19.1.2)$$

Теперь представим равенство (19.1.1) в следующем виде:

$$D(m, n) = D_1(m, n) - 2D_2(m, n) + D_3(m, n), \qquad (19.1.3)$$

где

$$D_1(m, n) = \sum_j \sum_k [F(j, k)]^2, \qquad (19.1.4а)$$

$$D_2(m, n) = \sum_j \sum_k F(j, k) T(j-m, k-n), \qquad (19.1.4б)$$

$$D_3(m, n) = \sum_j \sum_k [T(j-m, k-n)]^2. \qquad (19.1.4в)$$

Слагаемое $D_3(m, n)$ —это энергия эталона, которая постоянна и не зависит от координат (m, n). Энергия изображения в пределах окна, представленная первым слагаемым $D_1(m, n)$, при изменении

координат обычно меняется довольно медленно. Второе слагаемое — взаимная корреляция R_{FT} (m, n) изображения и эталона. При совпадении изображения и эталона взаимная корреляция должна быть велика, что приводит к малым значениям среднеквадратической ошибки. Однако величина взаимной корреляции не всегда адекватно отражает отличие изображения от эталона, поскольку энергия изображения D_1 (m, n) зависит от значений координат. Взаимная корреляция может увеличиться даже при отсутствии соответствия изображения эталону, если яркость изображения в окрестности точки с координатами (m, n) велика. Эту трудность можно обойти, сравнивая нормированную взаимную корреляцию

$$\tilde{R}_{FT}(m, n) = \frac{D_2(m, n)}{D_1(m, n)} = \frac{\sum_j \sum_k F(j, k)\, T(j-m, k-n)}{\sum_j \sum_k [F(j, k)]^2} \tag{19.1.5}$$

с порогом L_R (m, n). Считается, что сходство с эталоном имеет место, если

$$\tilde{R}_{FT}(m, n) > L_R(m, n). \tag{19.1.6}$$

Нормированная взаимная корреляция имеет максимальную величину, равную единице, тогда и только тогда, когда изображение в окне точно совпадает с эталоном. Главный недостаток метода сопоставления с эталоном заключается в необходимости использования огромного количества эталонов для учета изменений объектов, возникающих при их повороте и увеличении (уменьшении) размеров. По этой причине при сопоставлении с эталоном желательно ограничиться признаками, которые меньше зависят от изменений размера и формы объекта. Такими признаками могут быть, например, разветвления контурных линий, образующие фигуру типа Y или T.

19.2. СОГЛАСОВАННАЯ ФИЛЬТРАЦИЯ НЕПРЕРЫВНЫХ ДЕТЕРМИНИРОВАННЫХ ПОЛЕЙ

Согласованный фильтр, реализованный в виде электрической схемы, широко применяется для обнаружения одномерных сигналов, например, в радиолокации и связи [5—7]. Для обнаружения объектов на изображениях можно использовать двумерный согласованный фильтр [8—12].

Применяемый для обработки изображений согласованный фильтр представляет собой пространственный фильтр, выходной сигнал которого есть мера пространственной корреляции между входным и эталонным изображениями. Этот сигнал затем можно

использовать, например, для обнаружения заданного изображения или для пространственной привязки двух изображений.

Рассмотрим задачу обнаружения известного (эталонного) сигнала-изображения, описываемого функцией $F(x, y)$, на неизвестном входном изображении — функции $F_U\ (x, y)$, которое искажено воздействием аддитивного шума, т. е. стационарного случайного поля $N\ (x, y)$, независящего от $F\ (x, y)$. Таким образом, $F_U\ (x, y)$ состоит либо из сигнала-изображения и шума

$$F_U\ (x, y) = F\ (x, y) + N\ (x, y), \tag{19.2.1а}$$

либо только из шума

$$F_U\ (x, y) = N\ (x, y). \tag{19.2.1б}$$

Неизвестное изображение подвергается пространственной фильтрации согласованным фильтром с импульсным откликом $H\ (x, y)$ и частотной характеристикой $\mathscr{H}\ (\omega_x, \omega_y)$, на выходе которого образуется изображение, описываемое функцией

$$F_O(x, y) = F_U(x, y) * H\ (x, y). \tag{19.2.2}$$

Согласованный фильтр рассчитан так, чтобы отношение энергии сигнала-изображения к энергии шума достигало максимума в некоторой точке $(\mathscr{E}, \mathscr{N})$.

Мгновенное значение энергии сигнала-изображения в точке $(\mathscr{E}, \mathscr{N})$ на выходе фильтра при отсутствии шума есть

$$|S(\mathscr{E}, \mathscr{N})|^2 = |F\ (x, y) * H\ (x, y)|^2 \tag{19.2.3}$$

при $x = \mathscr{E}$ и $y = \mathscr{N}$. Согласно теореме о свертке,

$$|S\ (\mathscr{E}, \mathscr{N})\ |^2 = |\int\limits_{-\infty}^{\infty}\!\!\int \mathscr{F}\ (\omega_x, \omega_y)\ \mathscr{H}\ (\omega_x, \omega_y) \exp\{i\ (\omega_x\mathscr{E} + \omega_y\mathscr{N})\}\ d\omega_x\ d\omega_y\ |^2, \tag{19.2.4}$$

где $\mathscr{F}\ (\omega_x, \omega_y)$ — спектр Фурье сигнала $F\ (x, y)$. Предполагается, что аддитивный шум на входе $N\ (x, y)$ — стационарный, независимый от сигнала-изображения случайный процесс, который описывается спектральной плотностью $\mathscr{W}_N\ (\omega_x, \omega_y)$. Из формулы (1.10.8) следует, что полная мощность шума на выходе фильтра равна

$$N = \int\limits_{-\infty}^{\infty}\!\!\int \mathscr{W}_N\ (\omega_x, \omega_y)\ |\mathscr{H}\ (\omega_x, \omega_y)\ |^2\ d\omega_x\ d\omega_y. \tag{19.2.5}$$

Тогда для отношения сигнал/шум получаем

$$\frac{|S(\mathscr{E}, \mathscr{N})|^2}{N} =$$

$$= \frac{\left|\int\int_{-\infty}^{\infty} \mathscr{F}(\omega_x, \omega_y)\, \mathscr{H}(\omega_x, \omega_y) \exp\{i(\omega_x\mathscr{E} + \omega_y\mathscr{N})\}\, d\omega_x\, d\omega_y\right|^2}{\int\int_{-\infty}^{\infty} \mathscr{W}_N(\omega_x, \omega_y)\, |\mathscr{H}(\omega_x, \omega_y)|^2\, d\omega_x\, d\omega_y}. \quad (19.2.6)$$

Известно, что это отношение достигает максимума, если частотная характеристика фильтра имеет вид [5, 8]

$$\mathscr{H}(\omega_x, \omega_y) = \frac{\mathscr{F}^*(\omega_x, \omega_y) \exp\{-i(\omega_x\mathscr{E} + \omega_y\mathscr{N})\}}{\mathscr{W}_N(\omega_x, \omega_y)}. \quad (19.2.7)$$

Если входной шум является белым с равномерным спектром $\mathscr{W}_N(\omega_x, \omega_y) = n_W/2$, то частотная характеристика согласованного фильтра сводится к виду

$$\mathscr{H}(\omega_x, \omega_y) = \frac{2}{n_W} \mathscr{F}^*(\omega_x, \omega_y) \exp\{-i(\omega_x\mathscr{E} + \omega_y\mathscr{N})\} \quad (19.2.8)$$

и соответствующий импульсный отклик фильтра принимает вид

$$H(x, y) = \frac{2}{n_W} F^*(\mathscr{E} - x, \mathscr{N} - y). \quad (19.2.9)$$

В этом случае импульсный отклик согласованного фильтра пропорционален комплексно-сопряженной и повернутой на 180° функции, описывающей сигнал-изображение.

Для белого шума сигнал на выходе фильтра можно записать следующим образом:

$$F_O(x, y) = \frac{2}{n_W} F_U(x, y) * F^*(\mathscr{E} - x, \mathscr{N} - y), \quad (19.2.10\text{а})$$

или

$$F_O(x, y) = \int\int_{-\infty}^{\infty} \frac{2}{n_W} F_U(\alpha, \beta) F^*(\alpha + \mathscr{E} - x, \beta + \mathscr{N} - y)\, d\alpha\, d\beta. \quad (19.2.10\text{б})$$

Легко видеть, что если смещение согласованного фильтра $(\mathscr{E}, \mathscr{N})$ выбрано равным нулю, сигнал на выходе фильтра

$$F_O(x, y) = \frac{2}{n_W} \int\int_{-\infty}^{\infty} F_U(\alpha, \beta) F^*(\alpha - x, \beta - y)\, d\alpha\, d\beta \quad (19.2.11)$$

пропорционален коэффициенту корреляции функции входного изображения $F(x, y)$ и комплексно-сопряженного сигнала-изображения. Обычно параметры $\mathscr{E}$, $\mathscr{N}$ устанавливаются равными нулю, что соответствует отсутствию сдвига между эталонным сигналом и идентичной ему компонентой входного изображения.

Если неизвестное изображение состоит из сигнала-изображения, смещенного на расстояние Δx, Δy, и аддитивного шума:

$$F_U(x, y) = F(x + \Delta x, y + \Delta y) + N(x, y), \qquad (19.2.12)$$

то сигнал на выходе согласованного фильтра при $\mathscr{E} = 0$, $\mathscr{N} = 0$ будет равен

$$F_O(x, y) = \frac{2}{n_W} \int\int_{-\infty}^{\infty} [F(\alpha + \Delta x, \beta + \Delta y) + N(\alpha, \beta)] F^*(\alpha - x, \beta - y)\, d\alpha\, d\beta. \qquad (19.2.13)$$

Пик корреляции будет иметь место при $x = \Delta x$, $y = \Delta y$ на выходной плоскости; отсюда находится величина смещения входного изображения относительно эталонного. Следовательно, согласованный фильтр инвариантен относительно сдвига. Однако он не инвариантен относительно поворота изображения, которое требуется обнаружить.

В общем случае небелого шума можно осуществить согласованный фильтр, заданный уравнением (19.2.7), в виде двухступенчатого линейного фильтра с частотной характеристикой

$$\mathscr{H}(\omega_x, \omega_y) = \mathscr{H}_A(\omega_x, \omega_y)\, \mathscr{H}_B(\omega_x, \omega_y). \qquad (19.2.14)$$

Первая ступень, называемая отбеливающим фильтром, имеет частотную характеристику, выбранную так, что шум $N(x, y)$ со спектральной плотностью $\mathscr{W}_N(\omega_x, \omega_y)$ на его входе преобразуется в белый шум с единичной энергией на его выходе, так что

$$\mathscr{W}_N(\omega_x, \omega_y)\, |\mathscr{H}_A(\omega_x, \omega_y)|^2 \equiv 1. \qquad (19.2.15)$$

Частотную характеристику отбеливающего фильтра можно определить, представив спектральную плотность исходного шума в виде произведения [7]:

$$\mathscr{W}_N(\omega_x, \omega_y) = \mathscr{W}_N^+(\omega_x, \omega_y)\, \mathscr{W}_N^-(\omega_x, \omega_y). \qquad (19.2.16)$$

При этом должны выполняться следующие условия:

$$\mathscr{W}_N^+(\omega_x, \omega_y) = [\mathscr{W}_N^-(\omega_x, \omega_y)]^*, \qquad (19.2.17а)$$

$$\mathscr{W}_N^-(\omega_x, \omega_y) = [\mathscr{W}_N^+(\omega_x, \omega_y)]^*, \qquad (19.2.17б)$$

$$\mathscr{W}_N(\omega_x, \omega_y) = |\mathscr{W}_N^+(\omega_x, \omega_y)|^2 = |\mathscr{W}_N^-(\omega_x, \omega_y)|^2. \qquad (19.2.17в)$$

Можно положить, например,

$$\mathscr{W}_N^+(\omega_x, \omega_y) = \sqrt{\mathscr{W}_N(\omega_x, \omega_y)} \exp\{i\theta(\omega_x, \omega_y)\}, \quad (19.2.18)$$

где $\theta(\omega_x, \omega_y)$ — произвольный фазовый угол. В этом случае частотная характеристика отбеливающего фильтра равна

$$\mathscr{H}_A(\omega_x, \omega_y) = 1/\mathscr{W}_N^+(\omega_x, \omega_y). \quad (19.2.19)$$

На вход фильтра второй ступени поступает сигнал $F_1(x, y) + N_W(x, y)$, где $N_W(x, y)$ — белый шум с единичной спектральной плотностью и

$$F_1(x, y) = F(x, y) * H_A(x, y) \quad (19.2.20)$$

есть преобразованный сигнал-изображение со спектром

$$\mathscr{F}_1(\omega_x, \omega_y) = \mathscr{F}(\omega_x, \omega_y)\mathscr{H}_A(\omega_x, \omega_y) = \mathscr{F}(\omega_x, \omega_y)/\mathscr{W}_N^+(\omega_x, \omega_y). \quad (19.2.21)$$

Из формулы (19.2.8) находится оптимальная частотная характеристика фильтра второй ступени

$$\mathscr{H}_B(\omega_x, \omega_y) = \mathscr{F}_1^*(\omega_x, \omega_y) \exp\{-i(\omega_x\mathscr{E} + \omega_y\mathscr{N})\}, \quad (19.2.22а)$$

или

$$\mathscr{H}_B(\omega_x, \omega_y) = \frac{\mathscr{F}^*(\omega_x, \omega_y)}{\mathscr{W}_N^-(\omega_x, \omega_y)} \exp\{-i(\omega_x\mathscr{E} + \omega_y\mathscr{N})\}. \quad (19.2.22б)$$

Произведение $\mathscr{H}_A(\omega_x, \omega_y)\,\mathscr{H}_B(\omega_x, \omega_y)$ дает частотную характеристику оптимального фильтра, определяемую уравнением (19.2.7).

Основной недостаток согласованного фильтра, определяемого уравнением (19.2.7), состоит в том, что его частотная характеристика преимущественно зависит от энергии изображений, а не от их пространственной структуры. Для примера рассмотрим сигнал-изображение в виде яркого объекта шестиугольной формы, расположенного на черном фоне. Если неизвестное изображение содержит круглый диск такой же яркости и такой же площади, как и этот объект, то получающаяся в результате функция корреляции будет очень похожа на функцию корреляции, которая имеет место при идеальном сходстве. Вообще обычный согласованный фильтр дает относительно плохое разделение объектов различной формы, которые имеют почти одинаковые размеры и почти одинаковую энергию. Этот недостаток до некоторой степени преодолевается с помощью дифференциального согласованного фильтра [8], использующего контурную структуру объекта, который требуется обнаружить.

Частотная характеристика дифференциального согласованного фильтра p-го порядка задается выражением

$$\mathscr{H}_p(\omega_x, \omega_y) = \frac{(\omega_x^2 + \omega_y^2)^p \mathscr{F}^*(\omega_x, \omega_y) \exp\{-i(\omega_x \mathscr{E} + \omega_y \mathscr{N})\}}{\mathscr{W}_N(\omega_x, \omega_y)}, \qquad (19.2.23)$$

где p — целое. При $p = 0$ имеем обычный согласованный фильтр с

$$\mathscr{H}_0(\omega_x, \omega_y) = \frac{\mathscr{F}^*(\omega_x, \omega_y) \exp\{-i(\omega_x \mathscr{E} + \omega_y \mathscr{N})\}}{\mathscr{W}_N(\omega_x, \omega_y)}, \qquad (19.2.24)$$

а при $p = 1$ — фильтр, называемый согласованным «лапласовским» фильтром, частотная характеристика которого равна

$$\mathscr{H}_1(\omega_x, \omega_y) = (\omega_x^2 + \omega_y^2)\, \mathscr{H}_0(\omega_x, \omega_y). \qquad (19.2.25)$$

Его импульсный отклик равен

$$H_1(x, y) = \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right] * H_0(x, y). \qquad (19.2.26)$$

Частотную характеристику дифференциального согласованного фильтра p-го порядка можно записать следующим образом:

$$\mathscr{H}_p(\omega_x, \omega_y) = (\omega_x^2 + \omega_y^2)^p\, \mathscr{H}_0(\omega_x, \omega_y). \qquad (19.2.27)$$

Таким образом, дифференциальную согласованную фильтрацию можно выполнить, последовательно обрабатывая сигнал сначала обобщенным дифференциальным оператором, функция которого заключается в контрастировании перепадов на изображении, а затем обычным согласованным фильтром.

19.3. СОГЛАСОВАННАЯ ФИЛЬТРАЦИЯ ИЗОБРАЖЕНИЙ СЛУЧАЙНЫХ ПОЛЕЙ

В предыдущем разделе идеальное изображение, описываемое функцией $F(x, y)$, которое требовалось обнаружить в присутствии аддитивного шума, считалось детерминированным. Если состояние $F(x, y)$ известно не точно, а лишь с некоторой вероятностью, то идею согласованной фильтрации можно распространить на случай обнаружения случайного поля в присутствии шума [13]. Даже если известно, что функция $F(x, y)$ — детерминированная, часто оказывается полезным рассмотреть ее как случайную со средним $E\{F(x, y)\} = F(x, y)$. Такая формулировка позволяет учесть в процессе обнаружения априорные знания относительно пространственной корреляции элементов наблюдаемого изображения. При обычной согласованной фильтрации в соответствии с уравнением (19.2.7) эта корреляция полностью игнорируется.

Для целей анализа представим неизвестное поле либо в виде суммы идеального изображения $F(x, y)$ — реализации случайного поля (двумерного случайного процесса) с известными моментами — и шумового поля $N(x, y)$, не зависящего от изображения:

$$F_U(x, y) = F(x, y) + N(x, y), \tag{19.3.1а}$$

либо в виде одного лишь шума

$$F_U(x, y) = N(x, y). \tag{19.3.1б}$$

Неизвестное поле подвергается свертке с импульсным откликом согласованного фильтра $H(x, y)$; в результате на выходе фильтра образуется поле

$$F_O(x, y) = F_U(x, y) * H(x, y). \tag{19.3.2}$$

Согласованный фильтр для случайного изображения рассчитывается так, чтобы максимизировать отношение средней энергии сигнала (без учета шума) к дисперсии сигнала на выходе фильтра. Это простое обобщение обычного отношения сигнал-шум, описываемого формулой (19.2.6). При отсутствии шума энергия случайного сигнала в некоторой точке $(\mathscr{E}, \mathscr{N})$ выходного поля равна

$$|S(\mathscr{E}, \mathscr{N})|^2 = |E\{F(x, y) * H(x, y)\}|^2. \tag{19.3.3}$$

Согласно теореме о свертке и в силу линейности оператора математического ожидания,

$$|S(\mathscr{E}, \mathscr{N})|^2 = \left| \iint_{-\infty}^{\infty} E\{\mathscr{F}(\omega_x, \omega_y)\} \mathscr{H}(\omega_x, \omega_y) \exp\{i(\omega_x\mathscr{E} + \omega_y\mathscr{N})\} d\omega_x d\omega_y \right|^2. \tag{19.3.4}$$

Дисперсия на выходе согласованного фильтра в предположении стационарности и независимости сигнала и шума равна

$$N = \iint_{-\infty}^{\infty} [\mathscr{W}_F(\omega_x, \omega_y) + \mathscr{W}_N(\omega_x, \omega_y)] |\mathscr{H}(\omega_x, \omega_y)|^2 d\omega_x d\omega_y, \tag{19.3.5}$$

где $\mathscr{W}_F(\omega_x, \omega_y)$ и $\mathscr{W}_N(\omega_x, \omega_y)$ — спектральные плотности соответственно сигнала и шума. Обобщенное отношение сигнал/шум, совпадающее по форме с отношением (19.2.6) для детерминированного сигнала, максимизируется при

$$\mathscr{H}(\omega_x, \omega_y) = \frac{E\{\mathscr{F}^*(\omega_x, \omega_y)\} \exp\{-i(\omega_x\mathscr{E} + \omega_y\mathscr{N})\}}{\mathscr{W}_F(\omega_x, \omega_y) + \mathscr{W}_N(\omega_x, \omega_y)}. \tag{19.3.6}$$

Заметим, что если $F(x, y)$ — детерминированный сигнал, то выражение (19.3.6) сводится к выражению (19.2.7).

В выражении для частотной характеристики согласованного фильтра случайного сигнала часто заменяют среднее значение спектра идеального изображения, которое требуется обнаружить, на спектр самого изображения. В этом случае при $\mathscr{E} = \mathscr{N} = 0$

$$\mathscr{H}(\omega_x, \omega_y) = \frac{\mathscr{F}^*(\omega_x, \omega_y)}{\mathscr{W}_F(\omega_x, \omega_y) + \mathscr{W}_N(\omega_x, \omega_y)}. \qquad (19.3.7)$$

Интерес представляет особый случай, когда шум белый ($\mathscr{W}_N(\omega_x, \omega_y) = n_W/2$) и идеальное изображение рассматривается как неразделимый марковский процесс первого порядка, определяемый уравнением (1.9.17), энергетический спектр которого

$$\mathscr{W}_F(\omega_x, \omega_y) = 2/(\alpha^2 + \omega_x^2 + \omega_y^2), \qquad (19.3.8)$$

где $\exp(-\alpha)$ — коэффициент корреляции соседних элементов. Для таких процессов частотная характеристика согласованного фильтра имеет вид

$$\mathscr{H}(\omega_x, \omega_y) = \frac{2(\alpha^2 + \omega_x^2 + \omega_y^2)\,\mathscr{F}^*(\omega_x, \omega_y)}{4 + n_W(\alpha^2 + \omega_x^2 + \omega_y^2)}. \qquad (19.3.9)$$

На высоких пространственных частотах и при низком уровне шумов согласованный фильтр, определяемый соотношением (19.3.9), становится эквивалентным согласованному фильтру с лапласианом, который описывается выражением (19.2.25).

19.4. СОГЛАСОВАННАЯ ФИЛЬТРАЦИЯ ДИСКРЕТНЫХ ИЗОБРАЖЕНИЙ

Согласованный фильтр, предназначенный для обнаружения объектов, можно применять не только к непрерывным, но и к дискретным изображениям. Один из возможных подходов состоит в дискретизации непрерывной частотной характеристики согласованного фильтра (19.2.7) при помощи методов, описанных в разд. 11.4. Можно также разработать дискретный согласованный фильтр, используя векторное представление изображений [13, 14]. Такой подход, рассмотренный в настоящем разделе, представляется более целесообразным, так как он применим к нестационарным изображению и шуму. Кроме того, можно успешно бороться с краевыми эффектами, возникающими на границе изображений.

Рассмотрим наблюдаемое изображение в векторной форме, которое состоит из суммы вектора $\mathbf{f}$, представляющего детерминированное изображение, и вектора шума $\mathbf{n}$:

$$\mathbf{f}_U = \mathbf{f} + \mathbf{n} \qquad (19.4.1а)$$

или же только из шума

$$\mathbf{f}_U = \mathbf{n}. \qquad (19.4.1б)$$

Дискретная согласованная фильтрация осуществляется следующим образом: формируется скалярное произведение вектора $\mathbf{f}_U$ и вектора $\mathbf{m}$, описывающего импульсный отклик согласованного фильтра $\mathbf{m}$; в результате получается скаляр

$$f_O = \mathbf{m}^T \mathbf{f}_U. \tag{19.4.2}$$

Вектор $\mathbf{m}$ выбирается так, чтобы максимизировать отношение сигнал/шум. Мощность сигнала при отсутствии шума определяется как

$$S = (\mathbf{m}^T \mathbf{f})^2, \tag{19.4.3}$$

а мощность шума

$$N = E\{(\mathbf{m}^T \mathbf{n})(\mathbf{m}^T \mathbf{n})^T\} = \mathbf{m}^T \mathbf{K}_n \mathbf{m}, \tag{19.4.4}$$

где $\mathbf{K}_n$ — ковариационная матрица шума. Следовательно, отношение сигнал/шум равно

$$\text{С/Ш} = \mathbf{m}^T \mathbf{f} / \mathbf{m}^T \mathbf{K}_n \mathbf{m}. \tag{19.4.5}$$

Оптимальный вектор $\mathbf{m}$ можно определить, дифференцируя по $\mathbf{m}$ отношение сигнал/шум, заданное формулой (19.4.5), и приравнивая затем получающиеся производные нулю. Эти операции непосредственно приводят к соотношению

$$\mathbf{m} = \left[\frac{(\mathbf{m}^T \mathbf{K}_n \mathbf{m})}{(\mathbf{m}^T \mathbf{f})}\right] \mathbf{K}_n^{-1} \mathbf{f}, \tag{19.4.6}$$

где выражение в квадратных скобках — скаляр, который при соответствующей нормировке равняется единице. Таким образом,

$$f_O = \mathbf{f}^T \mathbf{K}_n^{-1} \mathbf{f}_U. \tag{19.4.7}$$

В случае белого шума действие согласованного фильтра сводится к определению скалярного произведения векторов $\mathbf{f}$ и $\mathbf{f}_U$. В общем случае ковариационную матрицу шума можно представить в виде произведения матриц

$$\mathbf{K}_n = \mathbf{K} \mathbf{K}^T, \tag{19.4.8}$$

где матрица $\mathbf{K} = \mathbf{E} \Lambda_n^{1/2}$, $\mathbf{E}$ — матрица, составленная из собственных векторов матрицы $\mathbf{K}_n$, а Λ_n — диагональная матрица соответствующих собственных значений [14]. Результат фильтрации можно рассматривать как скалярное произведение вектора $\mathbf{f}$ и неизвестного вектора $\mathbf{f}_U$, подвергнутых «отбеливанию» путем умножения на матрицу $\mathbf{K}^{-1}$:

$$f_O = [\mathbf{K}^{-1} \mathbf{f}]^T [\mathbf{K}^{-1} \mathbf{f}_U]. \tag{19.4.9}$$

Предыдущий вывод непосредственно распространяется на обнаружение случайных векторов. Выражение для энергии сигнала (19.4.3) принимает вид

$$S = (\mathbf{m}^T \mathbf{n}_f)^2, \qquad (19.4.10)$$

где $\mathbf{n}_f$ — вектор, компонентами которого являются средние значения компонент вектора $\mathbf{f}$, а дисперсия на выходе согласованного фильтра равна

$$N = \mathbf{m}^T \mathbf{K}_f \mathbf{m} + \mathbf{m}^T \mathbf{K}_n \mathbf{m} \qquad (19.4.11)$$

в предположении независимости $\mathbf{f}$ и $\mathbf{n}$. Получающееся в результате отношение сигнал/шум имеет максимальное значение при

$$\mathbf{m} = (\mathbf{K}_f + \mathbf{K}_n)^{-1} \mathbf{n}_f. \qquad (19.4.12)$$

Для получения согласованной фильтрации скалярное произведение векторов $\mathbf{m}$ и $\mathbf{f}_U$ можно определить непосредственно, используя соотношение (19.4.2) или же выражение (19.4.9), где $\mathbf{K} = \mathbf{E}\mathbf{\Lambda}^{1/2}$, а $\mathbf{E}$ и $\mathbf{\Lambda}$ — матрицы соответственно собственных векторов и собственных значений $(\mathbf{K}_f + \mathbf{K}_n)$ [14]. В частном, но достаточно распространенном случае белого шума и ковариационной матрицы разделимого марковского процесса первого порядка операцию «отбеливания» можно выполнить с помощью разработанного для винеровской фильтрации [15] эффективного алгоритма, в котором используется преобразование Фурье.

19.5. ПРИВЯЗКА ИЗОБРАЖЕНИЙ

Во многих прикладных задачах обработки изображений необходимо выполнять поэлементное сравнение двух изображений одного и того же объекта, зарегистрированных различными датчиками, или двух изображений некоторого объекта, полученных с помощью одного датчика, но в разное время. Чтобы осуществить такое сравнение, необходимо выполнить взаимную привязку этих изображений и таким образом скорректировать относительные пространственные сдвиги, различия в усилении, смещения, вызванные поворотом, а также геометрические искажения и искажения яркости каждого изображения. Часто оказывается возможным исключить или минимизировать воздействие большинства источников ошибок привязки путем соответствующей статической калибровки и коррекции датчиков. Однако в некоторых случаях обнаружение подобных ошибок и их последующую коррекцию требуется выполнять динамически для каждой пары изображений.

В настоящем разделе рассматривается задача совмещения изображений при наличии пространственного сдвига одного из них относительно другого. Полученные результаты можно использовать для компенсации различий, вызванных поворотом и изменением масштаба, если увеличить размерность задачи или вы-

полнить соответствующее преобразование координат. (Например при переходе к полярным координатам поворот приводится к сдвигу.)

19.5.1. КОРРЕЛЯЦИОННАЯ ПРИВЯЗКА

Классический способ взаимной привязки (совмещение) пары функций состоит в том, что формируется величина, измеряющая

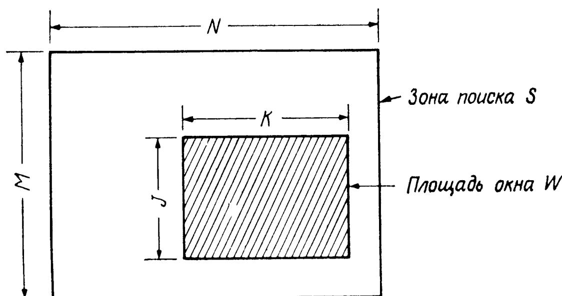


Рис. 19.5.1. Зона поиска при корреляционной привязке и площадь окна.

корреляцию между этими функциями, и находится положение максимума функции корреляции [16, 17]. Рассмотрим применение этого способа в случае двух измерений. Пусть массивы $F_1(j, k)$ и $F_2(j, k)$ представляют два дискретных изображения, которые требуется привязать друг к другу. В простейшей форме мера корреляции определяется следующим образом:

$$R(m, n) = \frac{\sum_{j=1}^{J} \sum_{k=1}^{K} F_1(j, k) F_2(j-m, k-n)}{\left[\sum_{j=1}^{J} \sum_{k=1}^{K} F_1^2(j, k)\right]^{1/2} \left[\sum_{j=1}^{J} \sum_{k=1}^{K} F_2^2(j-m, k-n)\right]^{1/2}}, \tag{19.5.1}$$

где (j, k) —индексы элементов в окне W размером $J \times K$ элементов, которое расположено внутри зоны поиска S размером $M \times N$ элементов. Рис. 19.5.1 иллюстрирует соотношение между зоной поиска и окном. Вообще, функцию корреляции требуется вычислить для всех $(M-J+1)(N-K+1)$ возможных смещений окна в зоне поиска для того, чтобы определить ее максимальное значение и получить оценку ошибки совмещения.

Применение этой простой меры корреляции связано с двумя основными трудностями. Во-первых, функция корреляции может иметь довольно размытый максимум, что затрудняет его обнару-

жение. Следует отметить, что мера корреляции (19.5.1) не учитывает пространственную структуру сравниваемых изображений. Во-вторых, шум на изображении может скрыть максимум корреляции. Обе трудности можно преодолеть, улучшив меру корреляции таким образом, чтобы в ней учитывались статистические свойства изображений $F_1(j, k)$ и $F_2(j, k)$.

Улучшенная мера корреляции определяется как

$$R_s(m, n) = \frac{\sum\limits_{j=1}^{J} \sum\limits_{k=1}^{K} G_1(j, k)\, G_2(j-m, k-n)}{\left[\sum\limits_{j=1} \sum\limits_{k=1}^{K} G_1^2(j, k)\right]^{1/2} \left[\sum\limits_{j=1}^{Z} \sum\limits_{k=1}^{K} G_2^2(j-m, k-n)\right]^{1/2}}, \tag{19.5.2}$$

где массивы $G_i(j, k)$ получаются сверткой массивов, описывающих изображения, с функциями $D_i(j, k)$, представляющими собой импульсные характеристики пространственных фильтров:

$$G_i(j, k) = F_i(j, k) * D_i(j, k). \tag{19.5.3}$$

Импульсные характеристики выбираются так, чтобы максимизировать пиковую корреляцию в том случае, когда сравниваемые изображения совмещены наилучшим образом. Эти характеристики можно получить с помощью теории согласованной фильтрации дискретных массивов, развитой в предыдущем разделе. Пусть $\mathbf{f}_1$ — вектор, полученный разверткой по столбцам фрагмента $F_1(j, k)$, соответствующего окну, а вектор $\mathbf{f}_2(m, n)$ составлен из элементов фрагмента $F_2(j, k)$ при заданном сдвиге (m, n). Полное число различных векторов $\mathbf{f}_2(m, n)$ составляет $M \cdot N$. Элементы векторов $\mathbf{f}_1$ и $\mathbf{f}_2(m, n)$ обычно сильно коррелированы. Поэтому в соответствии с методом согласованной фильтрации случайных полей первый шаг обработки должен состоять в «отбеливании», т. е. в умножении этих векторов на матрицы отбеливающих фильтров $\mathbf{H}_1$ и $\mathbf{H}_2$:

$$\mathbf{g}_1 = [\mathbf{H}_1]^{-1} \mathbf{f}_1, \tag{19.5.4а}$$

$$\mathbf{g}_2(m, n) = [\mathbf{H}_2]^{-1} \mathbf{f}_2(m, n). \tag{19.5.4б}$$

Матрицы $\mathbf{H}_1$ и $\mathbf{H}_2$ определяются через ковариационные матрицы изображений

$$\mathbf{K}_1 = \mathbf{H}_1 \mathbf{H}_1^T, \tag{19.5.5а}$$

$$\mathbf{K}_2 = \mathbf{H}_2 \mathbf{H}_2^T. \tag{19.5.5б}$$

Матрицы $\mathbf{H}_1$ и $\mathbf{H}_2$ можно представить в следующем виде:

$$\mathbf{H}_1 = \mathbf{E}_1 \mathbf{\Lambda}_1^{1/2}, \tag{19.5.6а}$$

$$\mathbf{H}_2 = \mathbf{E}_2 \mathbf{\Lambda}_2^{1/2}, \tag{19.5.6б}$$

где матрицы $\mathbf{E}_1$ и $\mathbf{E}_2$ образованы из собственных векторов ковариационных матриц $\mathbf{K}_1$ и $\mathbf{K}_2$, а $\mathbf{\Lambda}_1$ и $\mathbf{\Lambda}_2$ — диагональные матрицы из их собственных значений.

Мера корреляции (19.5.2) записывается в виде нормированного скалярного произведения

$$R_s(m, n) = \frac{g_1^T g_2(m, n)}{[g_1^T g_1]^{1/2} [g_2^T(m, n) g_2(m, n)]^{1/2}}. \qquad (19.5.7а)$$

Можно показать, что возможно другое представление этой меры:

$$R_s(m, n) = \frac{[(\mathbf{K}^T)^{-1} \mathbf{f}_1]^T \mathbf{f}_2}{\{[(\mathbf{K}^T)^{-1} \mathbf{f}_1]^T [(\mathbf{K}^T)^{-1} \mathbf{f}_1] [\mathbf{f}_2^T(m, n)] [\mathbf{f}_2(m, n)]\}}, \qquad (19.5.7б)$$

где $\mathbf{K} = \mathbf{H}_2\mathbf{H}_1^T$. Используя формулу (19.5.7*а*) надо производить «отбеливание» вектора $\mathbf{f}_1$ и всех $M \cdot N$ векторов $\mathbf{f}_2$, тогда как формула (19.5.7б) требует лишь одного умножения вектора $\mathbf{f}_1$ на матрицу $\mathbf{G} = (\mathbf{K}^T)^{-1}$. Ясно, что вторая формула предпочтительнее первой, если все вычисления выполняются обычными средствами.

Чтобы найти матрицы $\mathbf{G}$, необходимо вычислить два набора собственных векторов и собственных значений ковариационных матриц двух сравниваемых изображений в пределах окна. Если окно содержит $J \times K$ элементов, то каждая из ковариационных матриц $\mathbf{K}_1$ и $\mathbf{K}_2$ будет иметь размер $(J \cdot K) \times (J \cdot K)$. Например, если $J = K = 16$, то ковариационные матрицы $\mathbf{K}_1$ и $\mathbf{K}_2$ будут иметь размер 256×256. Вообще вычисление собственных векторов и собственных значений для таких больших матриц оказывается трудоемкой задачей для всех вычислительных машин, за исключением самых мощных. Однако в особых случаях эти вычисления можно заметно упростить [14]. Например, если изображения моделируются реализациями разделимого марковского процесса и отсутствует шум, то свертка (19.5.3) сводится к свертке изображения с маской обнаружения перепадов (17.4.6), когда

$$\mathbf{D}_i = \begin{bmatrix} \rho^2 & -\rho(1+\rho^2) & \rho^2 \\ -\rho(1+\rho^2) & (1+\rho^2)^2 & -\rho(1+\rho^2) \\ \rho^2 & -\rho(1+\rho^2) & \rho^2 \end{bmatrix}, \qquad (19.5.8)$$

где ρ — коэффициент корреляции смежных элементов изображения. Если оба изображения пространственно некоррелированы, то $\rho = 0$ и операция свертки не требуется. В другом предельном случае, когда $\rho = 1$,

$$\mathbf{D}_i = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}. \qquad (19.5.9)$$

Этот оператор представляет собой один из видов оператора Лапласа. Таким образом, когда изображения сильно коррелированы, вычисление улучшенной меры корреляции (19.5.2) сводится к обычной корреляции контурных изображений двух сцен.

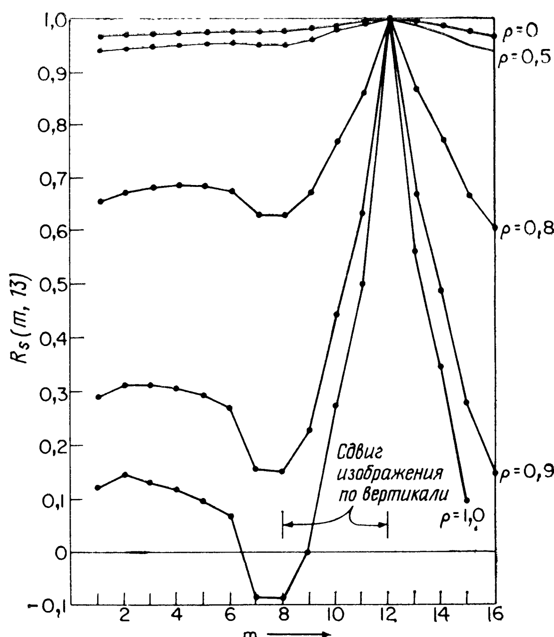


Рис. 19.5.2. Графики зависимости улучшенной меры корреляции изображений от величины сдвига.

На рис. 19.5.2 приведены результаты моделирования на вычислительной машине совмещения изображений естественных сцен с использованием улучшенной меры корреляции (19.5.7а). При моделировании изображение танка было смещено по горизонтали на три элемента и по вертикали на четыре элемента. Затем была вычислена мера корреляции этой пары изображений в окне размером 16×16 элементов для зоны поиска размером 32×32 элемента. На рис. 19.5.2 представлены кривые нормированной меры корреляции. Следует заметить, что при $\rho = 0$ (что соответствует обычной корреляции) довольно трудно различить пик величины $R_S(m, n)$. При $\rho = 0{,}9$ или больше $R_S(m, n)$ имеет четкий пик в точке, соответствующей правильному положению объекта.

19.5.2. ПОСЛЕДОВАТЕЛЬНАЯ ПРИВЯЗКА

Основной недостаток корреляционного способа привязки изображений состоит в том, что если площадь окна и зона поиска велики, то необходимо выполнить большой объем вычислений. С помощью корреляционного способа нельзя получить решение до тех пор, пока не будет вычислен корреляционный массив $R_S(m, n)$ для всех (m, n). Кроме того, объем вычислений для определения $R_S(m, n)$ одинаков при любом различии двух изображений. Эти недостатки привели к поиску последовательных алгоритмов, которые смогли бы по существу давать оценку величины различия изображений при меньшем числе вычислений.

Барнеа и Сильверман [18] предложили метод последовательных испытаний. Основная форма этого алгоритма обманчиво проста. Вычисляется мера различия изображений в окне

$$\mathscr{E}_S = \sum_j \sum_k |F_1(j, k) - F_2(j-m, k-n)|. \qquad (19.5.10)$$

Вычисления производятся последовательно от точки к точке. Если текущее значение меры различия превысит заранее определенный порог, прежде чем все $J \times N$ элементов в пределах окна будут исследованы, то считается, что испытание для данного окна не удалось, и проверяется новое окно. Если значение меры различия растет медленно, то число точек, рассмотренных до того момента, когда, наконец, порог будет превзойден, записывается и обозначается как оценка испытанного окна. После исследования всех окон окно с наибольшей величиной оценки считается находящимся в правильной позиции. Максимально возможная оценка равна числу точек в окне размером $J \times K$. Следует заметить, что в случае окон, которым соответствуют большие рассогласования, требуется относительно небольшое число вычислений.

Для увеличения скорости сходимости и повышения надежности предложено несколько модификаций алгоритма последовательных испытаний [18]. Естественной модификацией служит гибридная система, в которой для отбрасывания точек с большим рассогласованием используется алгоритм последовательных испытаний с последующим вычислением меры корреляции для оставшихся точек, подлежащих испытанию. Такая система могла бы сочетать преимущества рабочих характеристик способа измерения улучшенной меры корреляции и преимущества в скорости метода последовательных испытаний.

ЛИТЕРАТУРА

1. Duda R. O., Hart P. E., Pattern Classification and Scene Analysis, Wiley-Interscience, New York, 1973, pp. 276—284. [Имеется перевод: Дуда Р., Харт П., Распознавание образов и анализ сцен. — М.: Мир, 1976.]

2. Highleyman W. H., An Analog Method for Character Recognition, *IRE Trans. Electronic Computers*, **EC-10**, 3, 502—510 (September 1961).
3. Kanal L. N., Randall N. C., Recognition System Design by Statistical Analysis, Proceedings ACM National Conference, 1964.
4. Munson J. H., Experiments in the Recognition of Hand-Printed Text: Part I — Character Recognition, Proceedings Fall Joint Computer Conference, December 1968, pp. 1125—1138.
5. Turin G. L., An Introduction to Matched Filters, *IRE Trans. Inf. Theory*, **IT-6**, 3, 311—329 (June 1960).
6. Cook C. E., Bernfeld M., Radar Signals, Academic Press, New York, 1965. [Имеется перевод: Кук Ч., Бернфельд М., Радиолокационные сигналы. Теория и применение. — М.: Советское радио, 1971.]
7. Thomas J. B., An Introduction to Statistical Communication Theory, Wiley, New York, 1965, pp. 187—218.
8. Andrews H. C., Computer Techniques in Image Processing, Academic Press, New York, 1970, pp. 55—71. [Имеется перевод: Эндрюс Г., Применение вычислительных машин для обработки изображений. — М.: Энергия, 1977.]
9. Cutrona L. J., Leith E. N., Palermo C. J., Porcello L. J., Optical Data Processing and Filtering Systems, *IRE Trans. Inf. Theory*, **IT-6**, 3, 386—400 (June 1960).
10. Vander Lugt A., Rotz F. B., Kloester A., Jr., Character-Reading by Optical Spatial Filtering, in: Optical and Electro-Optical Information Processing, Tippett J. et al., Eds., MIT Press, Cambridge, Mass., 1965, pp. 125—141.
11. Vander Lugt A., Signal Detection by Complex Spatial Filtering, *IEEE Trans. Inf. Theory*, **IT-10**, 2, 139—145 (April 1964).
12. Kozma A., Kelly D. L., Spatial Filtering for Detection of Signals Submerged in Noise, *Appl. Opt.*, **4**, 4, 387—392 (April 1965).
13. Arcese A., Mengert P. H., Trombini E. W., Image Detection Through Bipolar Correlation, *IEEE Trans. Inf. Theory*, **IT-16**, 5, 534—541 (September 1970).
14. Pratt W. K., Correlation Techniques of Image Registration, *IEEE Trans. Aerospace and Electronic Systems*, **AES-10**, 3, 353—358 (May 1974).
15. Pratt W. K., Davarian F., Fast Computational Techniques for Pseudoinverse and Wiener Image Restoration, *IEEE Trans. Computers*, **C-26**, 6, 571—580 (June 1977).
16. Meyer-Eppler W., Darius G., Two-Dimensional Photographic Autocorrelation of Pictures and Alphabet Letters, in: Proceedings 3rd London Symposium on Information Theory, Cherry C., Eds., Academic Press, New York, 1956, pp. 34—36.
17. Anuta P. F., Digital Registration of Multispectral Video Imagery, *Soc. Photo-Optical Instr. Engineers J.*, **7**, 168—178 (September 1969).
18. Barnea D. I., Silverman H. F., A Class of Algorithms for Fast Image Registration, *IEEE Trans. Computers*, **C-21**, 2, 179—186 (February 1972).

Глава 20

СИСТЕМЫ ПОНИМАНИЯ ИЗОБРАЖЕНИЙ

Системы понимания изображений предназначены для анализа изображений, представленных функцией или массивом чисел, и составления описания изображенной сцены в «неизобразительной» форме. В простейшем виде система понимания изображений может просто сообщать о том, что на изображении имеется заданный объект или же, возможно, какой-нибудь неожиданный объект. В другом предельном случае такая система может создавать общее словесное описание сцены так же, как если бы человек написал о содержании данной картинки. Уже созданы системы, выполняющие относительно простые задачи описания изображений ограниченного класса. В настоящее время исследования направлены на развитие обобщенных программируемых систем, которые могли бы обрабатывать широкий класс изображений, возникающих в разнообразных задачах.

20.1. СИСТЕМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ

Многие задачи анализа изображений можно успешно решить с помощью классической системы распознавания образов, блок-схема которой показана на рис. 20.1.1 [1—4]. Системы распознавания образов предназначены для классификации входных изображений или их частей на несколько категорий. Например, применительно к задаче исследования земных ресурсов аэрофотографию полей, занятых под сельскохозяйственные культуры, можно разбить на маленькие фрагменты и каждый фрагмент классифицировать по типу выращиваемой культуры: пшеница, кукуруза, хлопок и т. д. Каждый фрагмент входного изображения можно рассматривать как точку в некотором пространстве всех возможных фрагментов. Для фрагмента размером $N \times N$ элементов, в котором каждое значение элемента квантовано на B уровней, число различных возможных фрагментов $B^{N \cdot N}$ может оказаться очень большим даже для относительно маленьких фрагментов и при грубом квантовании. Поэтому обычно возникает необходимость сократить это число посредством выделения признаков. Типичными признаками являются контурные точки, текстурные параметры и коэффициенты Фурье. Эти признаки также можно рас-

сматривать как точки в некотором пространстве. Для выработки алгоритма распознавания образов (классификации) измеряются признаки специально подобранных фрагментов изображений, точная классификация которых известна. Этот набор фрагментов

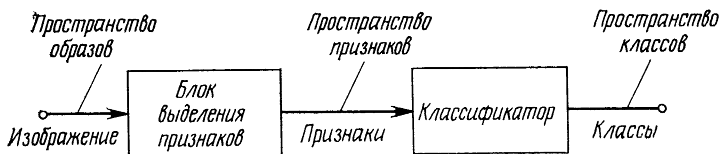


Рис. 20.1.1. Блок-схема системы распознавания образов.

называется обучающей выборкой. Результаты измерений признаков отображаются в пространство признаков. На рис. 20.1.2 показаны результаты измерения двух признаков (например, средней яркости и какого-либо параметра текстуры) фрагментов из обучающей выборки. Можно ожидать, что измеренные признаки обучающей выборки будут группироваться в соответствии с принадлежностью фрагментов к классам. В этом случае будут установлены границы областей принятия решений для разделения признаков новых фрагментов, подлежащих классификации. На ранних этапах развития распознавания образов многим представлялось, что это направление приведет к решению почти всех проблем, связанных с анализом изображений. Впоследствии этот оптимизм угас из-за двух ограничений методов распознавания образов. Во-первых, изображения имеют огромные размеры и во многих случаях число возможных классов очено велико. Поэтому обработка, которую должно производить устройство распознавания образов, часто оказывается невыполнимой. Более важное ограничение состоит в том, что методами классификации нельзя получить описание изображения, необходимое, например, для управления промышленным роботом. В качестве другого примера можно привести задачу структурного анализа сцены, для которой требуется получать описания типа «тело A распо-

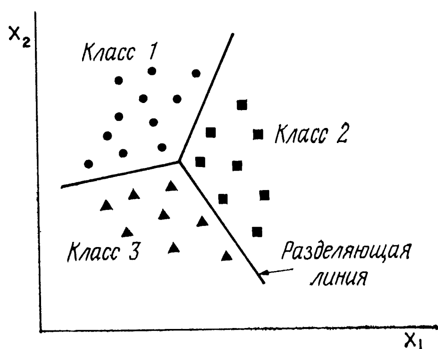


Рис. 20.1.2. Пример классификации обучающей выборки по двум признакам.

ложено выше и правее тела *В*» и т. д. Хотя во многих случаях методы классического распознавания образов оказываются непригодными для анализа изображений, от них не следует полностью отказываться. Существуют некоторые частные задачи анализа изображений, которые хорошо решаются методами распознавания образов.

20.2. МОДЕЛИ СИСТЕМ ПОНИМАНИЯ ИЗОБРАЖЕНИЙ

Для систем понимания изображений предложено множество разнообразных моделей. Исследование этих моделей показывает, что в их состав входит одинаковый набор блоков: блоки выделения признаков, символического представления и семантической интерпретации. В основном модели отличаются организацией этих блоков, характером управления ими, а также уровнем искусственного интеллекта и используемыми знаниями.

На рис. 20.2.1 представлена блок-схема иерархической модели системы понимания изображений, в которой используется «восходящая» процедура обработки. На первом этапе из наблюдаемого изображения выделяются простейшие признаки, такие, как яркость

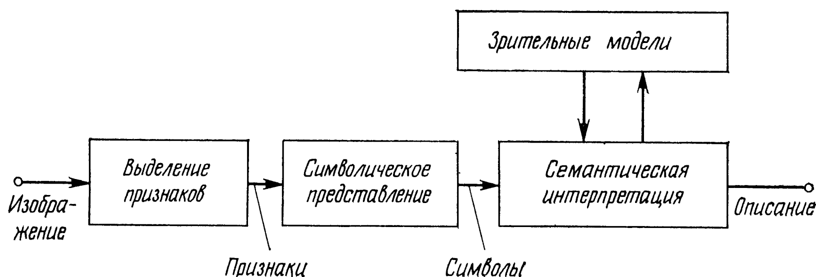


Рис. 20.2.1. Блок-схема иерархической модели системы понимания изображений при восходящей обработке.

элементов изображения, координаты контурных точек, параметры текстуры и т. д. Для уменьшения шума и влияния взаимных помех или для усиления признаков, которые в дальнейшем требуется выделить, в этот этап можно также включить поэлементную и пространственную обработку. Затем этот набор признаков поступает в блок символического представления, который формирует символы из признаков. Например, контурные точки группируются в отрезки линий или замкнутые кривые. Элементы с одинаковой яркостью или текстурой объединяются в области, однородные по этим свойствам. Множество символов затем обрабатывается блоком семантической интерпретации, в результате чего

получается некоторое желаемое описание сцены. Ясно, что характер описания сцены зависит от конкретной задачи. В некоторых случаях желательно, чтобы описание было просто указанием на наличие или отсутствие в сцене отыскиваемого объекта или же, возможно, на движение этого объекта вверх, вниз, направо или налево. В другом предельном случае для целей интерпретации фотографий это описание могло бы представлять собой подробное сообщение о содержании сцены. В блоке семантической интерпретации выявляются и размечаются структуры, представленные символами на входе блока. Для этого используются априорные знания о сцене или о классе рассматриваемых сцен. Эти значения вводятся в зрительные модели изображений, которые следует обработать. Знания о сцене могут варьироваться от очень простого и сразу приходящего в голову описания объектов, согласно которому, например, отыскиваемый объект должен быть ярче фона, до весьма сложных описаний сцены, которые включают сетевую структуру отношений, отображающих ограничения на взаимные связи между частями сцены. Блок семантической интерпретации должен исследовать все символы из множества символов изображения для определения их взаимных пространственных отношений и затем сопоставить выявленные отношения со зрительной моделью. Например, большие голубые области, обнаруженные в верхней части сцены, которые примыкают к большим зеленоватым областям, расположенным внизу сцены, должны быть интерпретированы соответственно как небо и трава. По завершении интерпретации различные части этой сцены размечаются, в результате чего образуется описание изображения.

Термин «восходящий», примененный к модели понимания изображений (рис. 20.2.1), показывает направление прохождения данных и сигналов управления. На каждом этапе независимо решается своя задача обработки или управления, и результат передается на следующий этап. Успех этого подхода обусловливается сокращением «размерности» задачи от этапа к этапу. Набор признаков должен быть меньше числа элементов входного изображения, а набор символов — меньше набора признаков. Это последовательное сокращение числа обрабатываемых параметров обязательно, так как относительная сложность обработки обычно возрастает на каждом последующем этапе. Иерархические системы, работающие по восходящему принципу, можно успешно развить для применения в ограниченных областях, в которых задача получения описания относительно проста и допустимый диапазон изменения входных изображений узок. Как только нарушается любое из этих двух условий, сразу же возникают трудности.

Другой большой класс систем понимания изображений основан на «нисходящей» иерархической модели, блок-схема которой

изображена на рис. 20.2.2. В этой модели структурный анализ, выполняемый на этапе интерпретации, проводится под управлением обучающих или тестовых описаний сцены. Согласованный фильтр, описанный в гл. 19, является наиболее распространенным примером системы нисходящей обработки. При согласованной фильтрации (сопоставления с эталоном) выполняется поиск в сцене конкретного объекта или конкретной структуры. Результатом поиска является указание на наличие или отсутствие этого объекта или некоторая вероятность его присутствия. Описания сцены

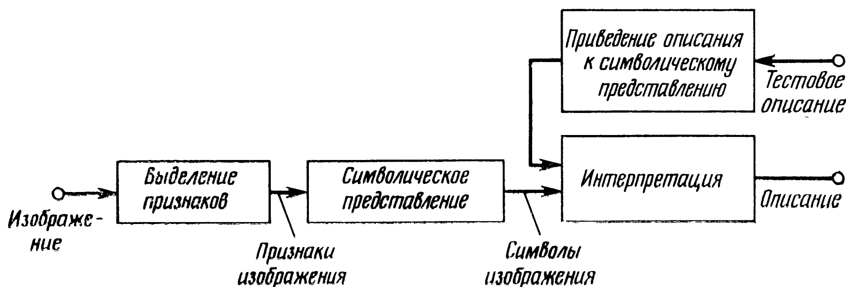


Рис. 20.2.2. Блок-схема иерархической модели системы понимания изображений при нисходящей обработке.

получаются путем последовательного поиска объектов в сцене, основанного на ранее достигнутых успехах или неудачах, и последующей интерпретации отношений между обнаруженными объектами. В примере согласованной фильтрации интерпретация обычно выполняется на уровне элементов изображения путем взаимной корреляции эталона объекта с наблюдаемым изображением. Интерпретация может выполняться на более высоком уровне путем корреляции признаков или символов, а не элементов изображения. Интерпретация на более высоком уровне потенциально имеет преимущества вследствие меньшей «размерности» задачи.

Иерархические системы понимания изображений, как правило, не обладают адаптивностью. Они обычно рассчитаны на решение специальных задач по анализу изображений ограниченного типа. Приспособление к другим задачам и другим типам изображений часто требует больших переделок систем. Другой недостаток иерархического подхода состоит в необходимости осуществления большого числа разнообразных обработок. Кроме того, многие из них выполняются напрасно, так как генерируются признаки и символы, которые не требуются для анализа данной сцены. Чтобы обойти эти трудности, в систему вводят центральное управляющее устройство, которое следит за работой всех

элементов системы и выдает команды различным блокам видоизменить свою работу с тем, чтобы оптимизировать работу всей системы и повысить ее эффективность. Минский и Пейперт [5, стр. 10] предложили другой гетерархический принцип организации систем понимания изображений, при котором эта система является целенаправленной, обладает зрительными знаниями и осуществляет распределенное управление. На рис. 20.2.3 приведена блок-схема такой системы, в которой с помощью распре-

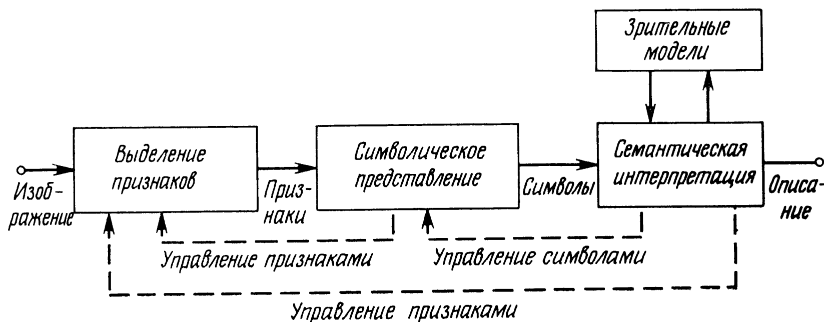


Рис. 20.2.3. Блок-схема гетерархической модели системы понимания изображений.

деленных обратных связей достигается гетерархическая организация работы. Если в данной модели на этапе семантической интерпретации возникают трудности при работе с набором входных символов, то по цепи обратной связи может быть передан сигнал в блок символического представления, чтобы видоизменить набор символов. Это воздействие в свою очередь может вызвать передачу команды по цепи обратной связи из блока символического управления в блок выделения признаков, которая требует изменения набора признаков. При необходимости возможна также непосредственная передача сигналов обратной связи из блока семантической интерпретации в блок выделения признаков. Кроме гибкости такая организация модели обеспечивает значительные дополнительные выгоды, позволяя использовать набор признаков и набор символов минимального размера, так как эти наборы могут быть перестроены по команде.

Редди и Ньюэлл [6] предложили другую модель системы понимания изображений, которая была названа моделью школьной доски. В этой модели, представленной в упрощенном виде на рис. 20.2.4, различные элементы системы сообщаются с общей памятью, т. е. школьной доской. Когда какой-нибудь элемент выполняет задачу, получающиеся результаты поступают в память, которая независимо доступна для всех других элементов,

Отдельные элементы могут либо управляться центральным устройством, либо обладать достаточным уровнем «интеллекта» и самостоятельно давать результаты, необходимые для выполнения общей задачи.

Наконец, следует отметить теорию «кадров» (frames), предложенную Минским [5, стр. 211, 7]. Минский выдвинул предположение, что, когда человек воспринимает обозреваемую им сцену, она соотносится с некоторым эталоном, который имеется в памяти

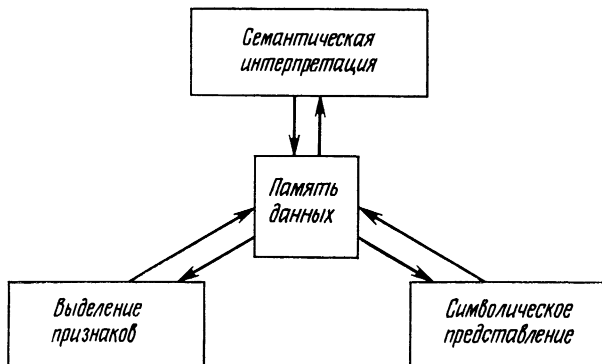


Рис. 20.2.4. Модель типа «школьная доска».

человека. Эти кадры являются в сущности моментальными снимками ранее виденных сцен. Человек легко может представить себе знакомые лица, места, события, например комнату, где он живет. В системе понимания изображений кадр — это некоторая структура, состоящая из сети узлов и отношений. Более высокие уровни этой сети фиксированы, так как служат «адресом», по которому кадр извлекается из памяти. Более низкие уровни сети содержат «пустоты», которые заполняются во время обзора сцены при условии, что справедливы определенные предположения. В примере с жилой комнатой можно ожидать, что диван будет расположен вдоль правой стены. Если это так, то воспринимаемые детали дивана помещаются в «пустоты» кадра. Если же диван оказался передвинутым к левой стороне комнаты, т. е. встретилась неожиданная ситуация, то имеющая отношение к положению дивана часть кадра, названная подкадром, должна быть изменена. Теория Минского в настоящее время не достигла уровня, необходимого для построения модели, однако она является, по-видимому, хорошей основой для будущего развития таких моделей.

Обсуждение представленных в этом разделе моделей явно поверхностно. Существующие в настоящее время модели имеют до некоторой степени более совершенную структуру, чем могло по-

казаться из приведенного изложения, но большинство моделей систем понимания изображений остается не законченными. В следующих разделах описаны конкретные подходы к разработке систем понимания изображений.

20.3. СИНТАКСИЧЕСКИЕ МЕТОДЫ

Школьники в начальных классах проводят много часов, разбивая предложения на существительные, местоимения, глаголы и т. д. Довольно неожиданно оказалось, что существует связь между этими скучными упражнениями и анализом сцен; методы формальной лингвистики, предназначенные для анализа естественного языка, были использованы для структурного и синтаксического анализа изображений [1]).

Формальные определения языка и его составных частей даны ниже.

Формальные определения языка

Синтаксис — описание языка.

Язык — множество предложений, составленных согласно некоторому набору правил.

Предложение — совокупность терминальных символов.

Терминальный символ — элемент предложения.

Нетерминальный символ — элемент правила подстановки (синтаксического правила).

Алфавит — некоторый конечный набор символов.

Грамматика — совокупность терминальных символов, нетерминальных символов, правил подстановки и начальных символов.

Правило подстановки (Production rule) — правило замены нетерминального символа другим нетерминальным символом или терминальным символом.

Грамматический разбор «сверху вниз» — последовательность применения правил подстановки, с помощью которых из нетерминального целевого символа строится предложение.

Грамматический разбор «снизу вверх» — последовательность применения правил подстановки, с помощью которых от предложения приходят к нетерминальному целевому символу.

В качестве примера рассмотрим простое предложение

$$S = \text{MIKE} \# \text{PAINTED} \# \text{THE} \# \text{PICTURE}, \qquad (20.3.1)$$

где $\#$ обозначает пробел. Для написания этого предложения используются следующие терминальные символы:

$$T = \{\text{MIKE, PAINTED, THE, PICTURE, } \#\}. \qquad (20.3.2)$$

---

[1]) Материал данного раздела составлен по учебному пособию Р. Х. Андерсона [8]. Работы [2, 9—12] также содержат полезные обзоры по этой теме.

Набор нетерминальных символов определяется следующим образом:

N = {⟨Предложение⟩, ⟨Группа существительного⟩, ⟨Глагольная группа⟩, ⟨Существительное⟩, ⟨Глагол⟩, ⟨Артикль⟩}. (20.3.3)

Нетерминальные символы участвуют в составлении предложений, но не входят в окончательный результат. Предложение строится с помощью набора правил подстановки. В данном случае они таковы:

⟨Предложение⟩ → ⟨Группа существительного⟩ # ⟨Глагольная группа⟩

⟨Группа существительного⟩ → ⟨Существительное⟩

⟨Группа существительного⟩ → ⟨Артикль⟩ # ⟨Существительное⟩

⟨Глагольная группа⟩ → ⟨Глагол⟩ || ⟨Группа существительного⟩

⟨Существительное⟩ → MIKE

⟨Существительное⟩ → PICTURE

⟨Глагол⟩ → PAINTED

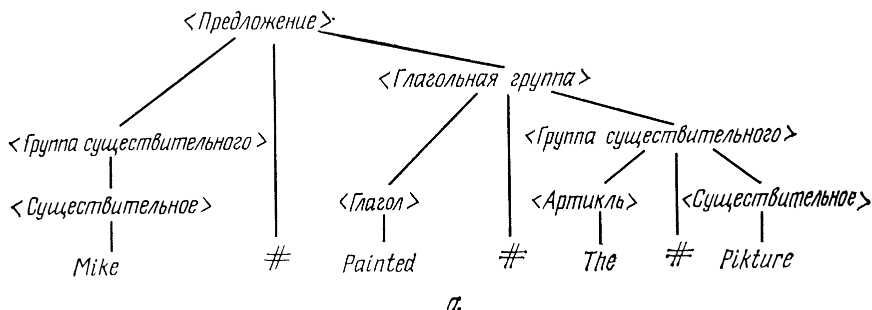
⟨Артикль⟩ → THE

Рис. 20.3.1, *а* иллюстрирует синтез предложения (20.3.1) посредством грамматического разбора «сверху вниз». На каждом этапе этой процедуры постепенно продвигаются от целевого символа ⟨Предложение⟩ к строке терминальных символов, заменяя правые части правила подстановки на их левые части. На конечном этапе строка найденных символов сравнивается с исходным предложением (20.3.1). Если эти предложения совпадают, то процедура заканчивается; в противном случае разбор должен начаться снова при другом порядке применения правил подстановки. Это продолжается до тех пор, пока не будет достигнуто совпадение или не будут исчерпаны все возможности. При использовании грамматического разбора «снизу вверх», схема которого показана на рис. 20.3.1, *б*, начинают с исходного предложения (20.3.1) и пытаются выполнить преобразования, которые могли бы привести к целевому символу ⟨Предложение⟩. Правильность преобразований контролируется правилами подстановки. Комбинации из правил подстановки должны проверяться до тех пор, пока не будет достигнут успех или же исчерпаны все возможности. Эффективность алгоритмов грамматического разбора исследована Гриффитсом и Петриком [13].

Синтаксический метод анализа цепочки букв алфавита для проверки ее правильности можно применить и к анализу изображений. При распространении синтаксического подхода на двумерные объекты возникают трудности. Во-первых, отсутствуют очевидные принципы выбора простых терминальных элементов. Что

представляет собой аналог буквы алфавита для картинки? Во-вторых, двумерность изображения порождает затруднения при разработке правил подстановки. Что представляет собой аналог цепочки слов для картинки?

На рис. 20.3.2 представлен пример грамматического разбора «сверху вниз» стилизованной сцены на терминальные макросим-



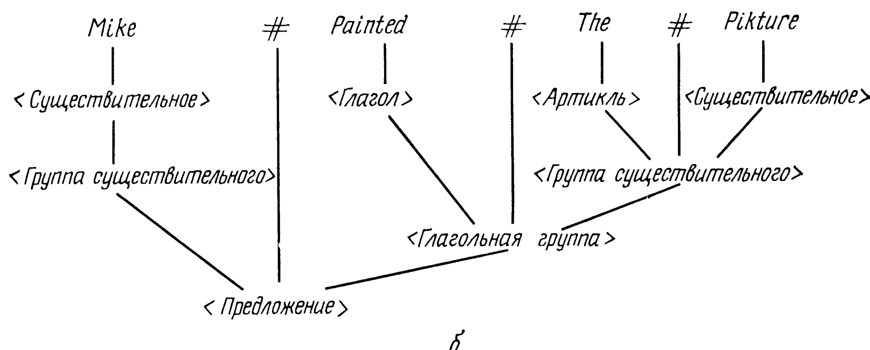
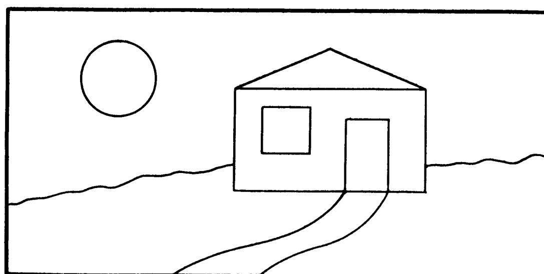


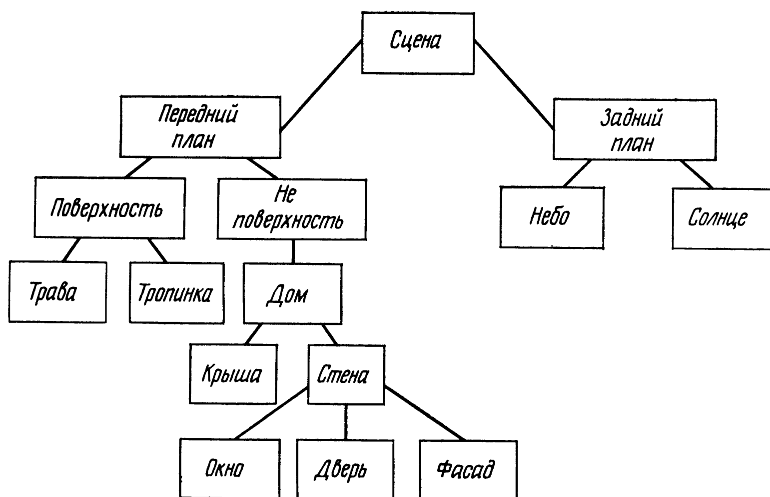
Рис. 20.3.1. Пример синтеза и анализа предложения.
а — синтез «сверху вниз»; *б* — анализ «снизу вверх».

волы: НЕБО, СОЛНЦЕ, ДВЕРЬ, КРЫША и т. д. Предположим, что эту сцену требуется проанализировать, чтобы определить, имеется ли на ней дом. Следуя по дереву разбора, дом определяется по треугольной крыше и стене, которая в свою очередь имеет окно и дверь. Исходя из общих представлений, дом можно распознавать путем поиска в сцене таких терминальных символов, как ОКНО, ДВЕРЬ, ФАСАД, и проверки их пространственных отношений, чтобы выяснить, образуют ли эти символы нетерминальный символ СТЕНА. Затем необходимо возобновить поиск, чтобы обнаружить символ КРЫША и проверить его струк-

турное отношение к символу СТЕНА. Если все правила подстановки выполняются, то считается, что дом в сцене имеется. Этот пример служит иллюстрацией трудностей, связанных с дву-



а



б

Рис. 20.3.2. Пример грамматического разбора сцены.
а — сцена; *б* — дерево разбора.

мерным синтаксическим анализом изображений. Успех этой процедуры основан на обнаружении довольно сложных терминальных макросимволов. Число макросимволов может быть достаточно большим при более широком классе сцен. Если макросимволы нельзя обнаружить с достаточной надежностью, то следует произвести разбор сцены при других более простых терминальных символах, таких, как отрезки линий и пятна. Это приводит к быстрому росту числа символов. С увеличением числа символов

растет набор правил подстановки. Значительные трудности при решении этих проблем ограничили возможность использования синтаксических методов анализом изображений, имеющих простую структуру. Впервые синтаксический анализ сцен был использован в работах [14] и [15] по символическому представлению рукописного текста. Нарасимхан [16, 17] впервые применил синтаксический метод к более общим классам изображений, уделив особое внимание обработке фотографий треков ядерных частиц в пузырьковых камерах. В модели Нарасимхана терминальные символы, названные основными множествами, состоят из отдельных отрезков линий, узлов, образованных этими линиями, и концов линий. Для анализа цепочек терминальных символов, обнаруженных на фотографии, применяется восходящий грамматический разбор. В соответствии с набором грамматических правил терминальные символы объединяются в цепочки нетерминальных символов, которые описывают такие части изображения, как прямые линии, незамкнутые кривые, изгибы U-образной формы и дуги окружностей. Эти цепочки затем объединяются в предложения.

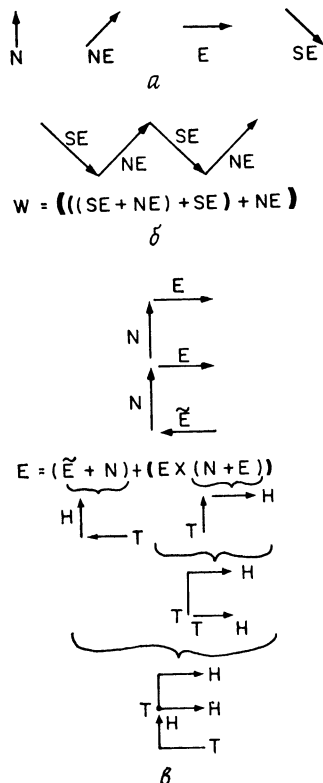


Рис. 20.3.3. Пример грамматического разбора изображений с помощью языка PDL. *а* — образцы линий; *б* — разбор буквы *W*; *в* — разбор буквы *E*.

Шоу [18, 19] разработал язык описания изображений PDL (Picture Description Language), основанный на синтаксических понятиях. В этом языке терминальные символы представлены отрезками линий, которые имеют метки «голова» и «хвост». Эти отрезки связываются простыми операторами:

$(A + B)$ голова A присоединена к хвосту B,
$(A \times B)$ хвост A присоединен к хвосту B,
$(A * B)$ голова A присоединена к голове B и хвост A присоединен к хвосту B,
$(\tilde{A})$ обозначение головы и хвоста в обратном (относительно A) порядке.

Составным структурам, которые порождаются этими операторами, также приписываются метки головы и хвоста. На рис. 20.3.3 приведен пример образования нескольких букв из простых прямолинейных отрезков. Шоу использовал язык PDL для грамматического разбора фотографий, полученных на искровых камерах.

Ледли [20, 21] применил синтаксические методы для машинной классификации хромосом. В его системе участки замкнутой границы хромосомы классифицируются на несколько категорий

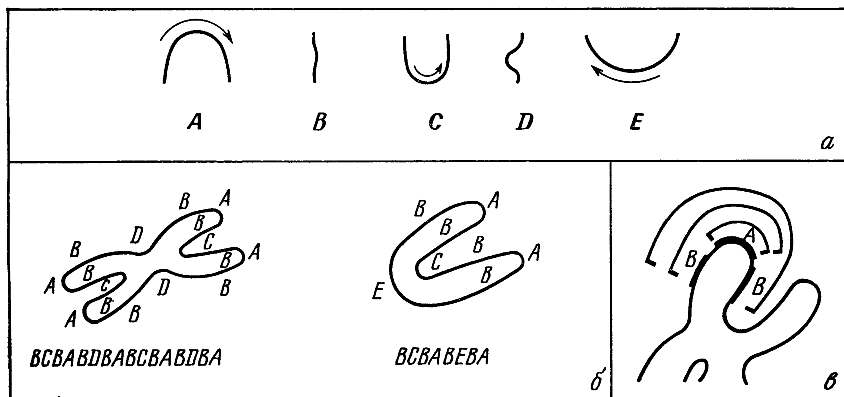


Рис. 20.3.4. Пример синтаксической классификации хромосом по Ледли [21]

отличающихся формой. На рис. 20.3.4 дан пример участков пяти таких категорий. В результате разбиения границы хромосомы на участки и классификации этих участков образуется строка терминальных символов, которая подвергается грамматическому разбору в соответствии с правилами подстановки. Затем полученное в результате разбора описание хромосомы сравнивается с описанием эталонных хромосом и производится классификация.

Исследования Нарасимхана, Шоу и Ледли показали, что синтаксические методы применимы к задачам анализа «простых» изображений. Вопрос о том, можно ли синтаксические приемы успешно распространить на классы более сложных изображений, остается открытым.

20.4. МЕТОДЫ, ИСПОЛЬЗУЮЩИЕ ОТНОШЕНИЯ МЕЖДУ ОБЪЕКТАМИ

Дерево грамматического разбора, описанное в предыдущем разделе, дает лишь грубое описание сцены с использованием связей между его терминальными и нетерминальными символами,

Часто бывает доступна и другая, более подробная информация. Например, может быть известно, что объект A расположен выше объекта B и что область C окружает оба объекта. Информацию такого типа можно представить в виде графа отношений, в котором ребра, соединяющие символы дерева разбора, размечены по типу

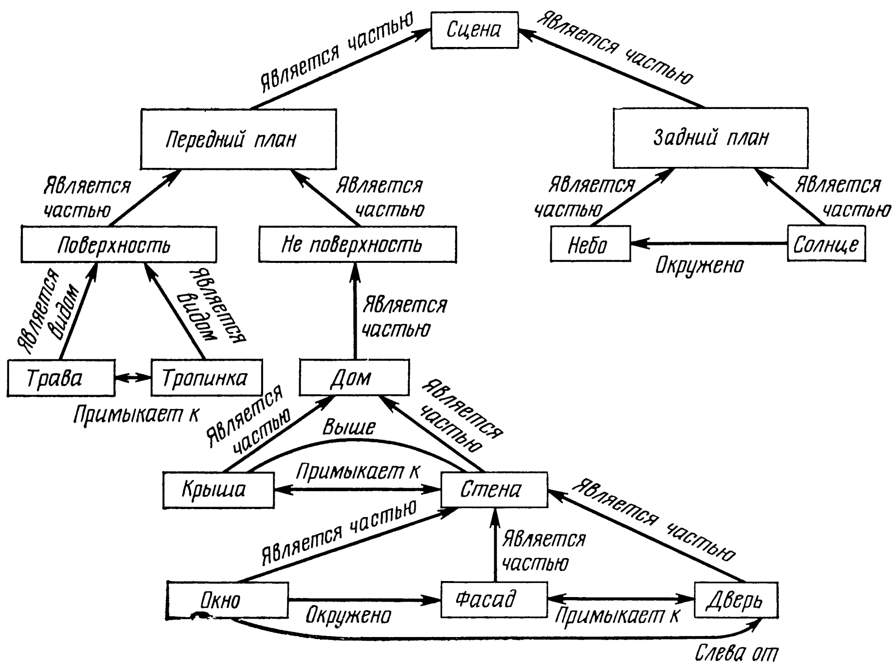


Рис. 20.4.1. Пример направленного графа отношений.

связи и пространственных отношений между символами. На рис. 20.4.1 приведен соответствующий пример. Стрелки на ребрах показывают направление отношения; например, КРЫША выше СТЕНЫ. Граф отношений с такими указателями называется направленным. Теоретический анализ методов, основанных на системах отношений, отсутствует. Остановимся на конкретных подходах к разработке систем понимания изображений, в которых явно или неявно используются эти отношения.

Впервые знания об отношениях между объектами были использованы для построения системы понимания изображений в работе Робертса [22], который разработал довольно полную систему программ анализа сцен, содержащих многогранники с плоскими поверхностями. Логично начать анализ сцен именно с многогранников, так как их можно относительно легко описать

и пространственные отношения между ними четко определены. В системе, предложенной Робертсом, анализ начинается с представления многогранника в виде контурного рисунка, который заносится в память в виде списка прямолинейных отрезков, заданных декартовыми координатами концевых точек. Этот список анализируется с целью обнаружения соединенных линейных отрезков, формирующих многоугольники. Рис. 20.4.2 иллюстрирует три типа многоугольников, которые могут находиться в сцене: выпуклый многоугольник, невыпуклый многоугольник и многоугольник, образованный границами изображения многогранника. На следующем этапе обработки множество вершин многоугольника сопоставляется с множеством вершин простых моделей объекта для того, чтобы разбить исходный объект на составные элементы, как показано в примере на рис. 20.4.3. Если найдено, что модель простого объекта является частью исходного объекта, то эта часть исключается из исходного объекта путем вычеркивания из списка и добавления соответствующих отрезков линий. Процесс продолжается до тех пор, пока не будет найдено соответствие исходного объекта одной из моделей простого объекта. Процесс сопоставления учитывает пространственные сдвиги, вращение и изменения масштаба моделей простых объектов по отношению к деталям исходного объекта.

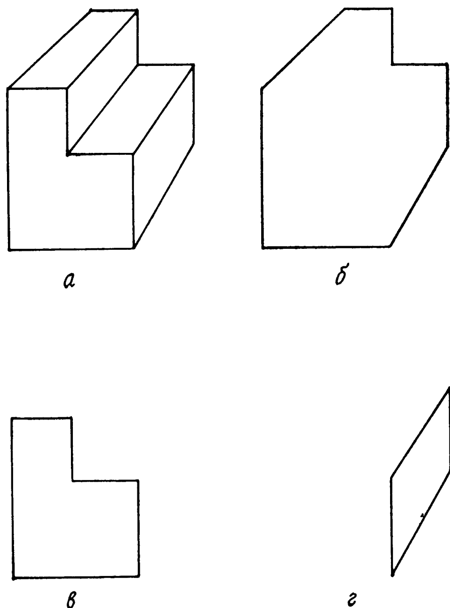


Рис. 20.4.2. Составление изображения многогранника из многоугольников.

а — объект; *б* — многоугольник, образованный внешним контуром; *в* — невыпуклый многоугольник; *г* — выпуклый многоугольник.

Идея Робертса относительно разбиения сложного многогранника на простые объекты породила направление исследовательских работ, которые теперь называются анализом блочных схем. Гузман [23] и Фальк [24] продолжили исследования сцен, содержащих «непрозрачные» объекты и представленных в виде неполного контурного рисунка. В предложенных ими системах выделение областей изображения, соответствующих плоским по-

верхностям сцены, осуществляется по специальным эвристическим правилам. Уинстон [5, стр. 5] предложил методы обучения, с помощью которых информация, полученная из анализируемых сцен, используется для наращивания набора зрительных моделей сцен. Модели запоминаются в виде направленного графа отношений, в котором дуги помечены простыми отношениями типа «соседство» и «опора» между объектами. На рис. 20.4.4 приведены направленный граф для сцены, состоящей из арки, и несколько примеров ее построения. Ключевым элементом работы Уинстона является понятие «почти то». В примере арки «почти то» может означать, что два объекта типа «опора» не могут быть смежными.

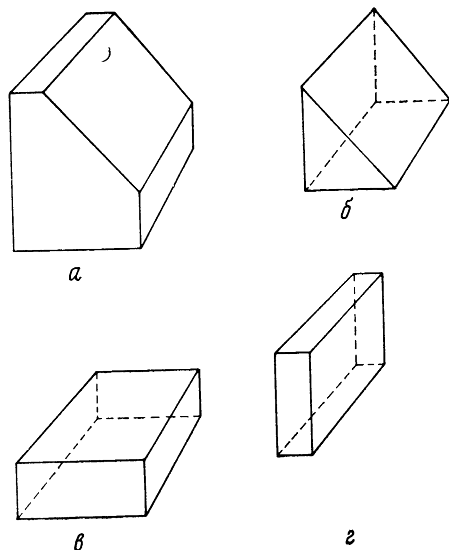
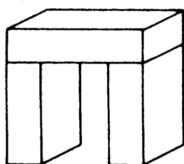


Рис. 20.4.3. Разборка многогранника.
а — объект; *б* — призма; *в*—основание в форме параллелепипеда; *г* — спинка в форме параллелепипеда.

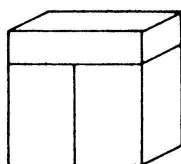
Представление многогранников в виде контурных рисунков не может быть произвольным; некоторые комбинации прямых линий и вершин не допускаются при контурном представлении многогранников. Клоуз [25] и Хаффман [26] разработали процедуры, позволяющие выяснить, реальные или нереальные (бессмысленные) объекты представлены на контурном рисунке. Уолц [5, стр. 19] продолжил разработку этого подхода, установив 11 категорий для разметки контурных линий (рис. 20.4.5). Если две или больше линий сходятся в некоторой вершине, то должны быть выполнены определенные ограничения. Например, в узле типа *L* градиенты тени, соответствующие двум линиям, должны быть одинаково направлены. Анализ всевозможных комбинаций разметок показал, что большая их часть физически не допустима. Полное число реальных типов вершин для многогранников составляет лишь несколько тысяч. В системе Уолца анализ начинается с разметки линий, выходящих из каждой вершины. Затем программа «фильтр» проверяет последовательность линейных меток связанных между собой вершин и исключает несовместимые метки. На следующем этапе с помощью набора комбинационных правил согласовываются неоднозначные метки. Их число в дальнейшем

а

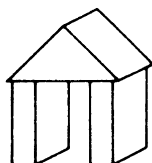
Арка



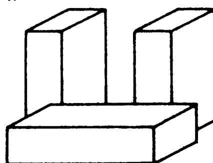
„Почти то"



Арка



„Почти то"



б

Рис. 20.4.4. Пример направленного графа отношений для арки [5].
а — направленный граф отношений; *б* — примеры.

может быть ограничено при анализе типов смежных вершин.

Указанные способы анализа блочных сцен обладают общим свойством: отправной точкой является либо абстрактный контурный рисунок, либо контурный рисунок, выделенный из реаль-

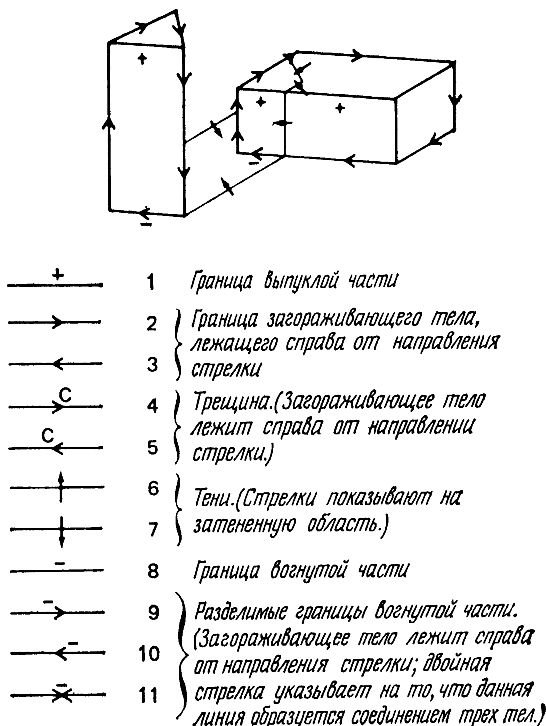


Рис. 20.4.5. Типы линий, образующих контуры многогранников [5].

ной сцены. Сираи [5, стр. 93, 27] развил другой подход к анализу сцен из многогранников, при котором реальные сцены анализируются непосредственно как массивы точек изображения. Программа Сираи — это первая программа, которая пытается выполнить гетерархический анализ осмысленных сцен. Знание типа объектов, с которыми она встречается, используется для управления поиском слабых перепадов яркости, соответствующих вогнутостям поверхностям объектов.

Важный шаг в направлении анализа сцен, содержащих тела с криволинейными поверхностями, сделан Хорном [5, стр. 115], а также Бинфордом, Эйджином и Невейшиа [28—30]. Хорн

использовал светотень освещенных объектов с гладкими криволинейными поверхностями для определения их физической формы. Работа Бинфорда, Эйджина и Невейшиа направлена на получение трехмерных описаний формы объектов; при этом были использованы срединноосные остовы и системы конических сечений объектов.

Одной из первых попыток анализа достаточно сложных реальных сцен с привлечением информации об отношениях между объектами была работа, выполненная Брайсом и Феннема [31]. В их системе первый этап обработки состоит в сегментации изображения по принципу наращивания областей. Затем граница каждой выделенной области аппроксимируется отрезками прямых. Далее на этапе анализа для группировки областей и распознавания объектов используются семантические знания. Например, в сцене «служебное помещение» семантические знания представляют собой следующие факты: пол находится в нижней части сцены, стена — на краю сцены, картина — на стене. Барроу и Поппльстоун [32] также разработали систему анализа сцен, основанную на сегментации метода наращивания областей. В их системе векторы признаков измеряются для каждой сегментированной области; при этом используются такие признаки, как форма, толщина, размер и отношение к соседним областям, а именно: выше, ниже, слева и т. д. Вектор признаков каждой области затем сопоставляется с наборами векторов признаков областей в зрительной модели класса сцен, подлежащих анализу. В некоторых случаях при сопоставлении возникают трудности, вызванные тенями и преградами.

Одна из трудностей, встречающихся непосредственно при анализе естественных сцен, состоит в том, что количество признаков изображения и символов слишком велико. Это множество данных вместе с огромным числом возможных комбинаций взаимных отношений переполняет структуры данных и делает бесполезным поиск по дереву.

Келли [33] сделал важное предложение: использовать планирование анализа сцен как способ ограничения размерности задач поиска по дереву. По схеме Келли признаки и символы выделяются сначала из изображения сцены, полученной с малым разрешением. Затем распознаются и размечаются объекты и вырабатывается гипотетическая модель сцены. Далее для выделения признаков и символов выполняется управляемый моделью поиск областей на картинке, полученной с полным разрешением.

Продолжая исследования Келли, Тененбаум и др. [34] предложили систему анализа, которая включает два этапа: ознакомление и обоснование. На этапе ознакомления исследуются простые характеристики изображения, основанные на зрительной

модели; при этом для управления поиском используется информация о содержании сцены. Например, в сцене служебного помещения поиск первоначально направлен на обнаружение плоской поверхности в средней части изображения, где предполагается наличие верхней поверхности письменного стола. Затем поиск продолжается в пределах этой поверхности для обнаружения интересующих нас объектов, которые анализируются и предварительно классифицируются. На этапе обоснования выполняется более тщательная обработка для выделения по возможности более надежных признаков в окрестности, где предполагается наличие объектов.

Полученная информация анализируется для определения правильности выдвинутой гипотезы. Таким образом ресурсы «дорогой» системы обработки информации можно распределить более эффективно.

Из приведенных материалов видно, что содержание данного раздела представляет собой просто краткий обзор результатов исследований в области построения систем понимания изображений.

В общем, достигнут некоторый прогресс, однако из большого числа интересных, но частных идей трудно выделить наиболее существенные. Ясно, что в этой важной области обработки изображений еще нужно проделать большую работу.

ЛИТЕРАТУРА

1. Andrews H. C., Introduction to Mathematical Techniques in Pattern Recognition, Wiley-Interscience, New York, 1972.
2. Duda R. O., Hart P. E., Pattern Classification and Scene Analysis, Wiley-Interscience, New York, 1973. [Имеется перевод: Дуда Р., Харт П., Распознавание образов и анализ сцен. — М.: Мир, 1976.]
3. Fukunaga K., Introduction to Statistical Pattern Recognition, Academic Press, New York, 1972. [Имеется перевод: Фукунага К., Введение в статистическую теорию распознавания образов. — М.: Наука, 1979.]
4. Meisel W. S., Computer-Oriented Approaches to Pattern Recognition, Academic Press, New York, 1972.
5. Winston P. H., The Psychology of Computer Vision, McGraw-Hill, New York, 1975. [Имеется перевод: Уинстон П., Психология машинного зрения. — М.: Мир, 1978.]
6. Reddy R., Newell A., Image Understanding: Potential Research Approaches, ARPA Image Understanding Workshop, Washington D. C., March 1975.
7. Winston P. H., Artificial Intelligence, Addison-Wesley, New York, 1977. [Имеется перевод: Уинстон П., Искусственный интеллект. — М.: Мир, 1980.]
8. Anderson R. H., Two-Dimensional Compiler Notes (неопубликованный курс), University of Southern California, July 1969.
9. Kirsch R. A., Computer Interpretation of English Text and Picture Patterns, *IEEE Trans. Electronic Computers*, **EC-13**, 4, 363—376 (August 1964).
10. Miller W. F., Shaw A. C., Linguistic Methods in Picture Processing — A Survey, Proceedings Fall Joint Computer Conference, December 1968, pp. 279—290.

11. Narasimhan R., On the Description, Generation, and Recognition of Classes in Pictures, in: Automatic Interpretation and Classification of Images, Graselli A., Ed., Academic Press, New York, 1969, pp. 1—42.
12. Fu K. S., Swain P. H., On Syntactic Pattern Recognition, in: Software Engineering, vol. 2, Tou J. T., Ed., Academic Press, New York, 1971, pp. 155—182.
13. Griffiths T. V., Petrick S. R., On the Relative Efficiencies of Context-Free Grammar Recognizers, *Commun. ACM*, **8**, 5, 289—300 (May 1965).
14. Grimsdale R. L., Automatic Pattern Recognition — New Morphological System Using a Digital Computer, *Wireless World*, **65**, 499—501 (November 1959).
15. Eden M., Handwriting and Pattern Recognition, *IRE Trans. Inf. Theory*, **IT-8**, 2, 160—166 (February 1962).
16. Narasimhan R., Labelling Schemata and Syntactic Description of Pictures, *Information and Control*, **7**, 151—179 (June 1964). [Имеется перевод: Нарасимхан Р. Метод маркировки и синтаксическое описание изображений. — В сб. Автоматический анализ сложных изображений. — М.: Мир, 1969, с. 87—111.]
17. Narasimhan R., Syntax-Directed Interpretation of Classes of Pictures, *Commun. ACM*, **9**, 3, 166—173 (March 1966). [Имеется перевод: Нарасимхан Р. Синтаксическая интерпретация классов изображений. — В сб. Автоматический анализ сложных изображений. — М.: Мир, 1969, с. 50—65.]
18. Shaw A. C., A Formal Picture Description Scheme as a Basis for Picture Processing Systems, *Information and Control*, **14**, 1, 9—52 (January 1969).
19. Shaw A. C., Parsing of Graph-Representable Pictures, *JACM*, **17**, 3, 453—481 (July 1970).
20. Ledley R. S., High-Speed Automatic Analysis of Biomedical Pictures, *Science*, **146**, 3641, 216—223 (October 1964).
21 Ledley R. S. et al., FIDAC: Film Input to Digital Automatic Computer and Associated Syntax-Directed Pattern-Recognition Programming System, in: Optical and Electro-Optical Information Processing, Tippett J. T. et al., Eds., MIT Press, Cambridge, Mass., 1965.
22. Roberts L. G., Machine Perception of Three-Dimensional Solids, in: Optical and Electro-Optical Information Processing, Tippett J. T. et al., Eds., MIT Press, Cambridge, Mass., pp. 159—197. [Имеется перевод: Робертс Л. Автоматическое восприятие трехмерных объектов. — В сб. Интегральные роботы. — М.: Мир, 1973, т. 1, с. 162—208.]
23. Guzman A., Decomposition of a Visual Scene into Three-Dimensional Bodies, Proceedings Fall Joint Computer Conference, Vol. 33, 1968, pp. 291—304. [Имеется перевод: Гузман А. Разбиение визуальной сцены на трехмерные тела. — В сб. Интегральные роботы. — М.: Мир, 1973, т. 1, с. 241—268.]
24. Falk G., Interpretation of Imperfect Line Data as a Three-Dimensional Scene, *Artificial Intelligence*, **3**, 2, 101—144 (1972).
25. Clowes M., On Seeing Things, Artificial Intelligence, **4**, 1, 79—116 (1971). [Имеется перевод: Клоувз М. Б. Как видеть вещи. — В сб. Интегральные роботы. — М.: Мир, 1975, т. 2, с. 89—135.]
26. Huffman D. A., Impossible Objects as Nonsense Sentences, in: Machine Intelligence, Vol. 6, Meltzer B., Mitchie D., Eds., University Press, Edinburgh, 1971, pp. 295—323.
27. Shirai Y., A Context Sensitive Line Finder for Recognition of Polyhedra, Artificial Intelligence, **4**, 2, 95—119 (1973).
28. Binford T. O., Agin G. J., Computer Description of Curved Objects, 3rd International Joint Conference on Artificial Intelligence, Menlo Park, Calif., 1973, pp. 629—640.
29. Nevatia R., Binford T. O., Structural Descriptions of Complex Objects, 3rd International Joint Conference on Artificial Intelligence, Menlo Park, Calif., 1973, pp. 641—657.

30. Nevatia R., Structured Descriptions of Complex Curved Objects for Recognition and Visual Memory, Springer-Verlag, New York, 1977.
31. Brice C. R., Fennema C. L., Scene Analysis Using Regions, Artificial Intelligence, 1, 205—206 (1970). [Имеется перевод: Брайс К. Р., Феннема К. Л. Анализ сцены при помощи выделения областей. — В сб. Интегральные роботы. — М.: Мир, 1975, т. 2, с. 136—159.]
32. Barrow H. G., Popplestone R. J., Relational Descriptions in Picture Processing, in: Machine Intelligence, Meltzer B., Mitchie D., Eds., University Press, Edinburgh, 1971, pp. 377—396.
33. Kelly M. D., Visual Identification of People by Computer, Stanford University Report AIM-130, 1970.
34. Tenenbaum J. M., Garvey J. M., Weyl S., Wolf H. C., An Interactive Facility for Scene Analysis Research, Stanford Research Institute, Technical Report 87, 1974.

Часть 6

Кодирование изображений

За последние 20 лет очень много усилий было направлено на разработку цифровых систем кодирования изображений для телевидения и фототелеграфии. Главная цель построения таких систем — сократить насколько возможно затрату двоичных единиц на кодирование изображения при сохранении приемлемой верности его воспроизведения. Экономное кодирование дает возможность: а) ускорить передачу отдельных изображений; б) увеличить число телевизионных программ, передаваемых одновременно по линии связи; в) снизить требования к мощности передатчика; г) уменьшить объем оборудования для запоминания изображений.

Как показывают психофизические исследования, скорость восприятия информации зрительной системой человека составляет около 50 дв. ед./с. С другой стороны, теоретико-информационный анализ изображений естественного происхождения указывает на их исключительно высокую избыточность. Тем не менее обычные телевизионные системы ежесекундно обрушивают на наблюдателя около 50 млн. дв. ед. Отсюда становится ясно, насколько велики потенциальные возможности совершенствования кодирующих систем.

Главы 21, 22 и 23 посвящены описанию различных методов кодирования. В гл. 24 рассматриваются вопросы эффективности систем кодирования изображений.

Глава 21

АНАЛОГОВЫЕ МЕТОДЫ КОДИРОВАНИЯ ИЗОБРАЖЕНИЙ

В этой главе дается классификация методов кодирования изображений и описывается процесс передачи изображения в вещательных телевизионных системах. Рассматриваются методы кодирования изображений, связанные с передачей непрерывных или дискретизованных сигналов.

21.1. КЛАССИФИКАЦИЯ МЕТОДОВ КОДИРОВАНИЯ ИЗОБРАЖЕНИЙ

Классификация цифровых систем кодирования изображений может быть проведена различными способами (табл. 21.1.1).

Таблица 21.1.1

Классификация систем кодирования изображений

Категория изображаемых объектов
Сцены естественного происхождения
Документация, визуализированные данные
Фотометрическое содержание изображений
Одноцветные изображения
Цветные изображения
Спектрозональные изображения
Шкала уровней яркости
Двухградационные изображения
Многоградационные изображения
Категория средств обработки сигналов
Аналоговые системы обработки
Цифровые системы обработки
Метод обработки сигналов
Построчная (одномерная) обработка
Внутрикадровая (двумерная) обработка
Межкадровая (трехмерная) обработка

Возможна классификация в зависимости от категории изображений, взятых в качестве объекта кодирования. С одной стороны, это могут быть изображения постоянно окружающих наблюдателя сюжетов и сцен естественного происхождения. С другой стороны, изображение может представлять собой результат двумерной визуализации данных, с которым человек сталкивается

лишь в специальных обстоятельствах; примером может служить радиолокационная диаграмма в координатах дальность—скорость. Другая классификация изображений основана на их фотометрических характеристиках (одноцветные, цветные и спектрозональные изображения). Ряд кодирующих систем предназначен для обработки только двухградационных изображений, например чертежей; другие же способны оперировать также многоградационными изображениями. Совершенно другой способ клас-

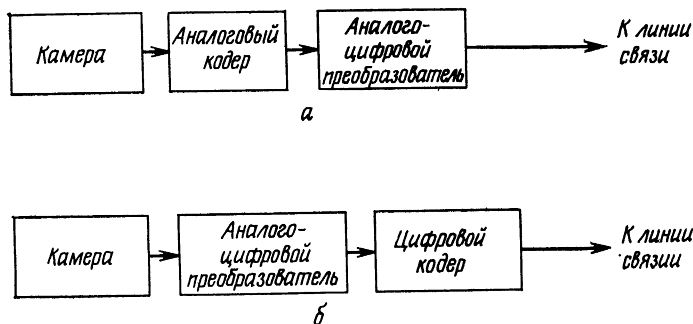


Рис. 21.1.1. Сжатие изображений посредством кодирования.
а — аналоговое кодирование; *б* — цифровое кодирование.

сификации иллюстрирует рис. 21.1.1. Здесь определяющим признаком является метод обработки сигналов — аналоговый или цифровой. Развертка изображения дает на выходе видеодатчика непрерывный электрический сигнал. В аналоговой системе этот сигнал подвергается обработке аналоговыми средствами с целью сокращения полосы его частот, что дает возможность снизить частоту дискретизации и передавать меньшее количество отсчетов. В цифровой системе выходной сигнал видеодатчика прежде всего подвергается аналого-цифровому преобразованию, а затем цифровой обработке, в ходе которой производится перекодирование для сокращения количества передаваемых двоичных единиц. Наконец, кодирующие системы различаются тем, что повторяющийся процесс обработки может в один прием охватывать отдельную строку развертки, некоторую часть плоскости изображения или при передаче подвижных изображений, в том числе в реальном масштабе времени, определенное число последовательных кадров.

В последующих главах рассматриваются операции по обработке сигналов в системах кодирования изображений. Исчерпывающая информация об этих системах содержится в работах [1—7].

21.2. СИСТЕМЫ ВЕЩАТЕЛЬНОГО ТЕЛЕВИДЕНИЯ

Передача одноцветного изображения по проводной линии связи была впервые продемонстрирована независимо друг от друга Бэйрдом и Дженкинсом в 1925 г. [8]. В 1928 г. Бэйрд впервые продемонстрировал цветную телевизионную систему с последовательной передачей красного, зеленого и синего полей изображения. В последующие годы Айвс и Джонсрад [9] осуществили телевизионную систему с одновременной передачей цветов. Бэйрду принадлежит заслуга проведения первой вещательной телевизионной передачи по трансатлантической линии радиосвязи в 1928 г. Цветное телевизионное вещание впервые было осуществлено Голдмарком в 1940 г. [10, 11]. Со времени этих ранних экспериментов предложено и испытано большое число методов передачи изображений и телевизионных систем. В настоящее время всемирное распространение получили три стандарта вещательного телевидения: системы НТСЦ (Национального комитета по телевизионным системам США), ПАЛ (система с построчной коммутацией фазы) и СЕКАМ (система с последовательной передачей цветов и запоминанием) [12]. На Североамериканском континенте и в Японии применяется система НТСЦ. Система СЕКАМ принята во Франции, в восточноевропейских странах и в СССР. Остальные европейские страны, включая Великобританию, а также южноамериканские и африканские страны, используют систему ПАЛ.

21.2.1. СИСТЕМА НТСЦ [13—15]

Национальный комитет по телевизионным системам (НТСЦ) США был образован в 1940 г. с целью выработки стандартов для систем вещательного одноцветного телевидения. НТСЦ был сформирован вновь в 1950 г. с тем, чтобы изучить предложенные к тому времени системы цветного телевидения и дать рекомендации по выбору стандартной системы для вещательного телевидения США.

В одноцветном варианте системы НТСЦ электрический сигнал, описывающий передаваемую сцену, создается на выходе телевизионной камеры в результате развертки по строкам. Полный цикл развертки включает последовательное образование четного и нечетного полей, составленных соответственно четными и нечетными строками. Телевизионным кадром называют совокупность обоих полей; кадр содержит в общей сложности 525 строк. На формирование каждого поля отводится $^1/_{60}$ с, а кадра в целом — $^1/_{30}$ с. Выбор частоты кадров определяется частотой переменного тока в энергосетях США, составляющей 60 Гц. Если бы частота полей не совпадала с этой величиной, то взаимодействие электро-

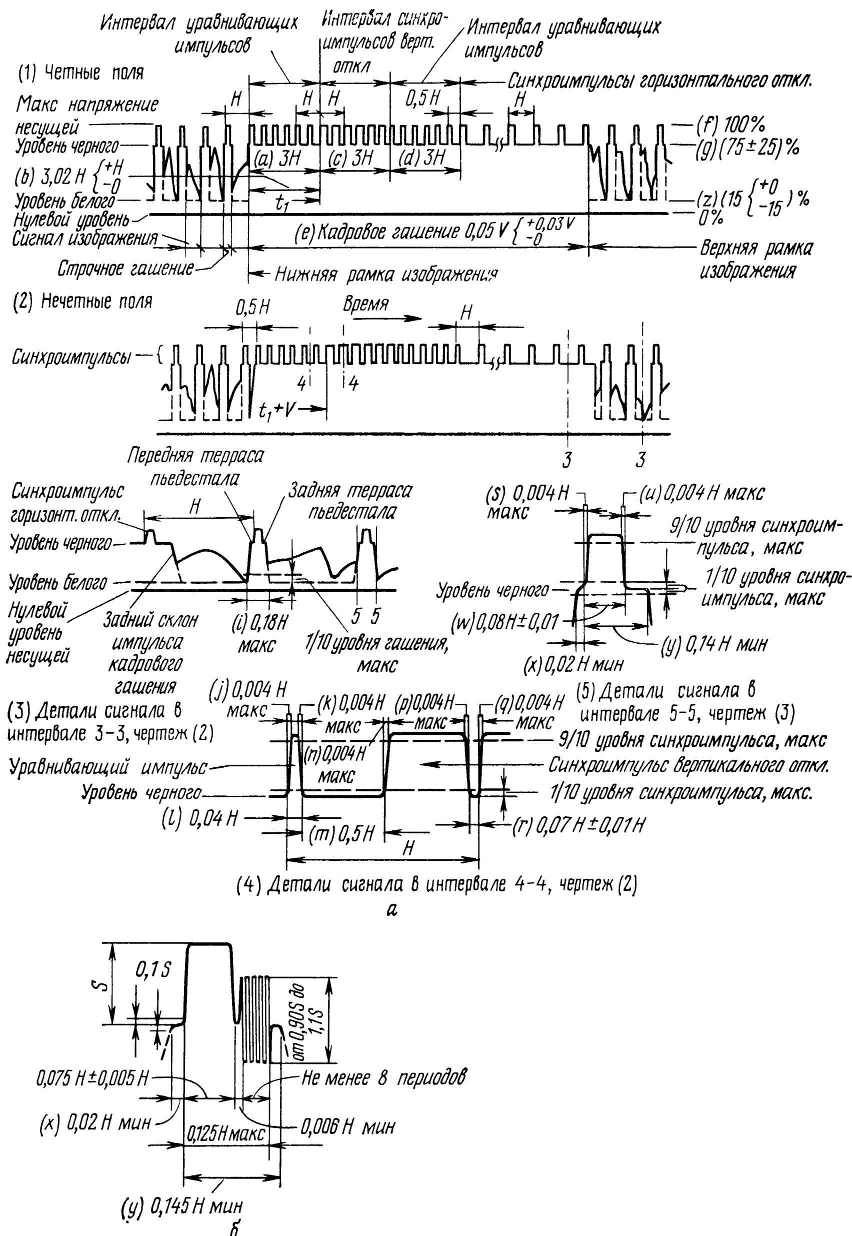


Рис. 21.2.1. Стандартный сигнал синхронизации системы НТСЦ [15].
а — сигналы синхронизации при передаче одноцветного изображения; *б* — детали сигнала синхронизации при передаче цветного изображения в интервале 5—5, чертеж (3).

магнитных полей, создаваемых генераторами развертки и устройствами электропитания, привело бы к помехам на изображении за счет наводок на комбинационных частотах как в передатчике, так и в приемнике. По окончании каждой строки передается синхроимпульс горизонтальной развертки; передача каждого поля завершается синхроимпульсом вертикальной развертки. На рис. 21.2.1 показаны типичные видеосигналы. Перед тем как использовать полный (т. е. содержащий синхроимпульсы) аналоговый видеосигнал для модуляции несущей, его пропускают через фильтр нижних частот с полосой пропускания около 4 МГц; коэффициент передачи по мощности на верхней граничной частоте составляет 0,5. В приемнике после демодуляции производится выделение синхроимпульсов из полного видеосигнала для управления цепями горизонтального и вертикального отклонений телевизионного видеомонитора.

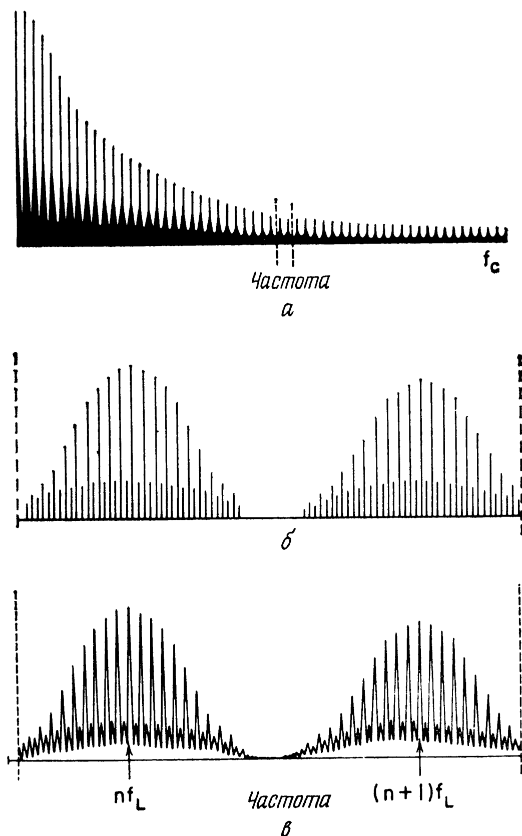


Рис. 21.2.2. Спектр сигнала яркости (воспроизводится с некоторыми упрощениями по работе [15]).

а — спектр сигнала яркости; *б* — спектр при передаче белого поля — увеличенный участок спектра *а*; *в* — спектр при передаче типичной сцены — увеличенный участок спектра *а*.

В связи с периодичностью строчной и кадровой разверток спектр видиосигнала имеет дискретный характер. Это обстоятельство иллюстрируется рис. 21.2.2. Спектр здесь несколько идеализирован и соответствует передаче неподвижного изображения [15, стр. 8—65]. Линии спектра образуют ряд сгущений, располагающихся вдоль оси частот с одинаковым интервалом, равным 15 750 Гц (частота строчной развертки). Как показано на рис. 21.2.2, *б*, каждое сгущение в свою очередь представляет собой последовательность спектральных линий, разделенных интервалом 60 Гц (частота полей). Скачки в амплитуде от одной линии к соседней — следствие чересстрочной развертки. Наличие в передаваемой сцене подвижных объектов ведет к определенному расширению линий спектра (рис. 21.2.2, *в*).

В цветной телевизионной системе процесс развертки по строкам сопровождается формированием на выходе камеры трех сигналов, соответствующих красной, зеленой и синей составляющим передаваемой сцены. Комбинация этих трех сигналов

$$Y = 0{,}299R + 0{,}587G + 0{,}114B \qquad (21.2.1)$$

определяет яркость изображения. Сигнал яркости Y передается затем совместно с сигналами цветности I и Q, представляющими собой линейные функции цветоразностных сигналов $R—Y$ и $B—Y$. Цветоразностный сигнал $G—Y$ может быть получен как

$$G - Y = -0{,}581\,(R - Y) - 0{,}394\,(B - Y), \qquad (21\ 2.2)$$

а затем уже могут быть восстановлены сигналы R, G и B сложением яркости Y с каждым из цветоразностных сигналов. Цветоразностные сигналы подвергаются нормировке согласно соотношениям

$$U = (R - Y)/1{,}14 = 0{,}877\,(R - Y), \qquad (21.2.3а)$$

$$V = (B - Y)/2{,}03 = 0{,}493\,(B - Y), \qquad (21.2.3б)$$

принятым в качестве стандарта в системах НТСЦ, ПАЛ и СЕКАМ. Нормировка проводится так, чтобы при переходе от одноцветной к цветной телевизионной передаче размах полного видеосигнала повышался не более чем в заданном отношении 1,33 : 1. В системе НТСЦ сигналы цветности I и Q формируются с помощью соотношений

$$I = U\cos(33°) - V\sin(33°), \qquad (21.2.4а)$$

$$Q = U\sin(33°) + V\cos(33°). \qquad (21.2.4б)$$

В то же время между сигналами Y, I и Q, с одной стороны, и выходными сигналами камеры R, G, B, с другой стороны, может быть установлена непосредственная связь:

$$Y = 0{,}299R + 0{,}587G + 0{,}114B, \tag{21.2.5а}$$

$$I = 0{,}596R - 0{,}274G - 0{,}322B, \tag{21.2.5б}$$

$$Q = 0{,}211R + 0{,}523G + 0{,}312B. \tag{21.2.5в}$$

Процесс обработки и кодирования видеосигнала в передатчике системы НТСЦ иллюстрируется схемой на рис. 21.2.3. Прежде всего производится фильтрация низких частот: для спектра сигнала яркости Y устанавливается полоса частот около 4 МГц, а для спектров сигналов I и Q — соответственно 1,3 и 0,5 МГц. Достаточно жесткое ограничение полосы частот сигналов I и Q не проходит, разумеется, бесследно для восстановленных на

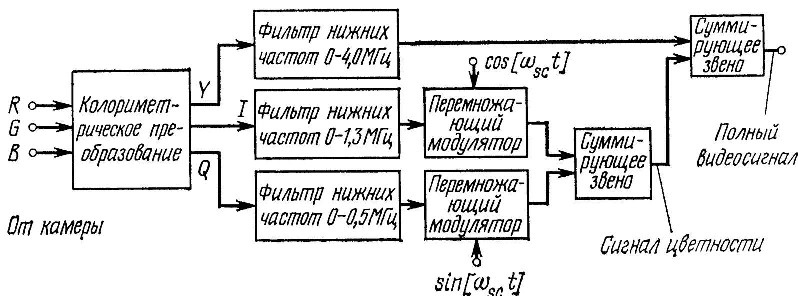


Рис. 21.2.3. Процесс кодирования цветного изображения в системе НТСЦ; нормирование цветного видеосигнала.

приемной стороне сигналов цвета R, G, B. Однако это ограничение практически не отражается на визуальном восприятии изображения, так как по отношению к цветовым составляющим изображения пространственная разрешающая способность зрительной системы человека сравнительно невелика [16]. Далее уровень каждого из сигналов I и Q модулирует по амплитуде соответствующую поднесущую, причем между двумя поднесущими установлен фазовый сдвиг, равный 90°. Сложение модулированных поднесущих дает объединенный сигнал цветности, который складывается с сигналом яркости. На рис. 21.2.4 показаны огибающие спектров сигнала яркости и объединенного сигнала цветности.

Как отмечалось выше, спектр видеосигнала в случае телевизионного растра имеет линейчатую структуру, а его энергия рассредоточена между отдельными линиями вдоль всей отведенной полосы частот. Таким образом, при формировании объединенного сигнала цветности его спектр следует расположить на оси частот так, чтобы сложение сигналов яркости и цветности не приводило к наложению линий их спектров. Это достигается с помощью частотного сдвига, кратного половине частоты строчной развертки; спектры обоих сигналов располагаются при этом подобно паре гребенок, сдвинутых так, что зубья одной из них приходятся на просветы между зубьями другой. Таким способом сигналы яркости и цветности могут быть объединены в цветовой видеосигнал, полоса частот которого находится в тех же пределах, какие установлены для обычных систем одноцветного телевидения. Наконец, в интервалы цветового видеосигнала, соответствующие обратному ходу луча, вводятся синхроимпульсы строчной и кадровой разверток, что дает полный цветовой видеосигнал. В приемнике в результате отделения синхроимпульсов произ-

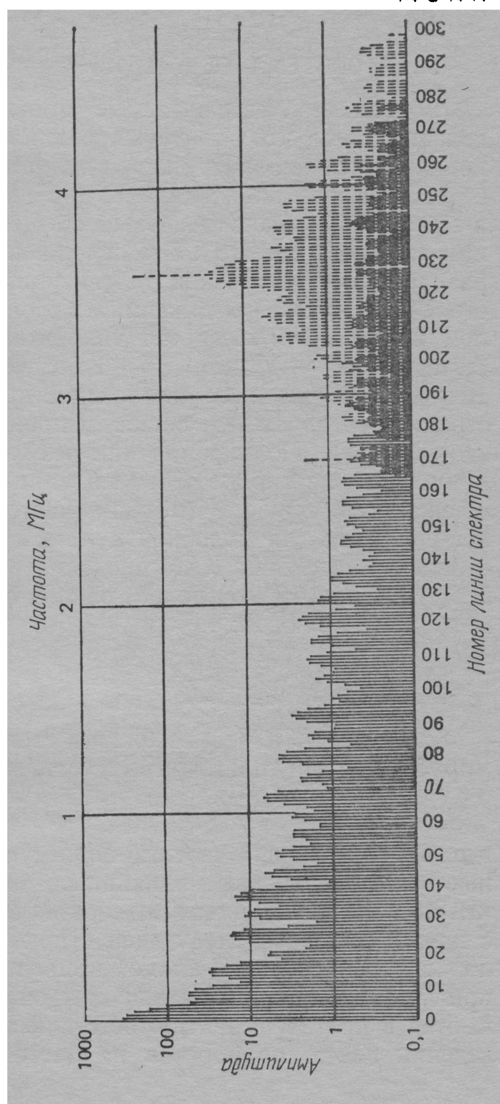


Рис. 21.2.4. Сложение спектров сигнала яркости и сигнала цветности [15].

водится восстановление цветового видеосигнала. Затем демодулятор разделяет сигналы Y, I, Q, результат преобразования которых, согласно соотношениям

$$R = 1{,}000Y + 0{,}956I + 0{,}621Q, \qquad (21.2.6а)$$

$$G = 1{,}000Y - 0{,}272I - 0{,}647Q, \qquad (21.2.6б)$$

$$B = 1{,}000Y - 1{,}106I + 1{,}703Q \qquad (21.2.6в)$$

используется для управления видеомонитором.

21.2.2. СИСТЕМА СЕКАМ [17]

Возражения, выдвигаемые против системы НТСЦ, обычно связаны с тем, что для формирования компоненты цветности здесь используется квадратурная модуляция. Квадратурная модуляция ведет к фазовым искажениям при выделении сигнала цветности, а тем самым и к значительным смещениям цвета при восстановлении изображения [18]. Это осложнение послужило причиной для разработки французской телевизионной системы СЕКАМ. В системе СЕКАМ на протяжении каждой из строк совместно с сигналом яркости Y передается какой-либо один из цветоразностных сигналов, $R - Y$ или $B - Y$, с чересстрочным чередованием. Декодер, запоминая цветоразностный сигнал на протяжении строки, повторяет этот сигнал в процессе воспроизведения следующей строки. Получаемая как результат чередования одиночная компонента цветности путем модуляции поднесущей объединяется с сигналом яркости методом частотного перемежения, какой используется для подобной цели в системе НТСЦ. Французский стандарт системы СЕКАМ предусматривает четкость по вертикали 625 строк при частоте полей, равной 50 Гц, с коэффициентом перемежения строк 2 : 1.

21.2.3. СИСТЕМА ПАЛ

Система ПАЛ, весьма близко стоящая к системе НТСЦ, формирует полный цветовой видеосигнал как

$$C = Y + U \sin \omega_{SC}t + kV \cos \omega_{SC}t, \qquad (21.2.7)$$

где ω_{SC} — угловая частота поднесущей, U и V — нормированные цветоразностные сигналы, определяемые соотношением (21.2.3), а коэффициент k составляет по модулю единицу, но с началом каждой следующей строки меняет знак на противоположный. Чередование знаков приводит к тому, что возникающие при передаче искажения для каждой пары соседних строк оказываются в противофазе. Соответствующие погрешности в воспроизведении цвета, приобретая на цветовом графике противополож-

ные направления, в процессе визуального восприятия в значительной степени компенсируют друг друга. В пределах Европы принят стандарт системы ПАЛ с четкостью 625 строк при частоте полей 50 Гц и с коэффициентом перемежения строк 2 : 1. В некоторых южноамериканских странах используется система ПАЛ с частотой полей 60 Гц в соответствии с частотой переменного тока в местных энергосетях.

21.3. МОДУЛЯЦИОННЫЕ МЕТОДЫ СОКРАЩЕНИЯ ПОЛОСЫ ЧАСТОТ

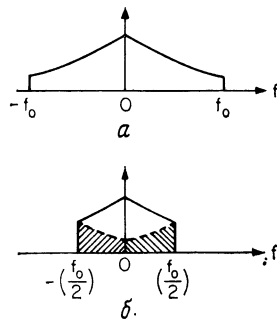


Рис. 21.3.1. Метод частотного перемежения.

а — спектр исходного видеосигнала; *б* — «складывание» спектра.

Уже давно было замечено, что обычная телевизионная развертка дает сигнал, спектр которого имеет достаточно четко выраженную дискретную структуру: на интервалы между отдельными компонентами спектра приходится весьма незначительная часть энергии. Это обстоятельство послужило основой для разработки экспериментальной одноцветной телевизионной системы с сокращенной полосой передаваемых частот [20]. Использованный в этой системе метод сжатия спектра получил название частотного перемежения: как показано на рис. 21.3.1, спектр как бы складывается вдвое путем поворота его высокочастотной половины вокруг границы с низкочастотной половиной, так чтобы компоненты обеих половин перемежались вдоль оси частот. Эта операция может быть выполнена с помощью обычных методов модуляции и фильтрации; для разделения обеих половин спектра при восстановлении сигнала требуются гребенчатые фильтры. Процедура восстановления сигнала должна проводиться с особой тщательностью, чтобы не было низкочастотного «ползания» строк при воспроизведении изображения. В принципе система с частотным перемежением достаточно проста в осуществлении, однако коэффициент сжатия спектра при передаче аналоговыми средствами не может превышать в такой системе 2 : 1.

21.4. УМЕНЬШЕНИЕ РАЗРЕШАЮЩЕЙ СПОСОБНОСТИ И ЧАСТОТЫ КАДРОВ

Одним из простейших способов сокращения полосы передаваемых частот является ограничение пространственной разрешающей способности видеодатчика и частоты повторения кадров. Однако вопрос о том, насколько можно уменьшать разрешающую

способность и частоту кадров, во многих случаях представляет значительные трудности в связи с необходимостью оценки субъективного восприятия качества изображения или фотометрической точности его воспроизведения.

21.4.1. УМЕНЬШЕНИЕ ПРОСТРАНСТВЕННОЙ РАЗРЕШАЮЩЕЙ СПОСОБНОСТИ

Минимальную величину пространственной разрешающей способности системы обработки или передачи изображений легко вычислить, если к ней предъявлены чисто фотометрические требования, определяющие угловую разрешающую способность в пределах заданного поля зрения. Если же изображения, воспроизводимые такой системой, приходится оценивать тем, в какой мере они поддаются визуальному анализу и распознаванию или обеспечивают возможность обнаружения каких-либо объектов, то выработка требований к разрешающей способности значительно усложняется. Руководством в подобных обстоятельствах может до некоторой степени служить гл. 7, где рассматриваются различные показатели качества изображений.

Как отмечалось в разд. 21.2, для вещательных телевизионных систем различных стран существуют различия в стандартах четкости изображения. В системах цифровой обработки изображений количество строк часто определяется какой-либо целочисленной степенью двух (256, 512, 1024, ...), чем обеспечивается полное использование цифровых регистров. При этом создается возможность наиболее эффективного применения быстрых вычислительных алгоритмов, как, например, быстрого преобразования Фурье.

Во многих случаях нет практической необходимости в высокой разрешающей способности по всей площади изображения. Иногда заранее известно, какой именно район поля зрения представляет наибольший интерес, в других же случаях наиболее важные участки изображения могут быть выделены оператором. Майер и Гарденхайр [21] разработали метод передачи фотоснимков с переменной разрешающей способностью. При подготовке фотоснимка к передаче оператор очерчивает участки, требующие кодирования с высокой или низкой разрешающей способностью. Роль участков, кодируемых с низким пространственным разрешением, состоит в воспроизведении общей обстановки, окружающей наиболее значащие объекты. Примеры кодирования этим методом, полученные путем цифрового моделирования, показаны на рис. 21.4.1. Следует отметить, что кодирование с переменной разрешающей способностью требует передачи сведений не только об уровнях элементов, но и о границах между участками изображения. Как показало моделирование, для многих приложений

коэффициент сокращения полосы частот может достигать значения 8 : 1.

Переменная пространственная разрешающая способность может быть использована также для сокращения полосы частот в тех случаях, когда объект, представляющий главный интерес для получателя изображений, расположен в центральной зоне поля

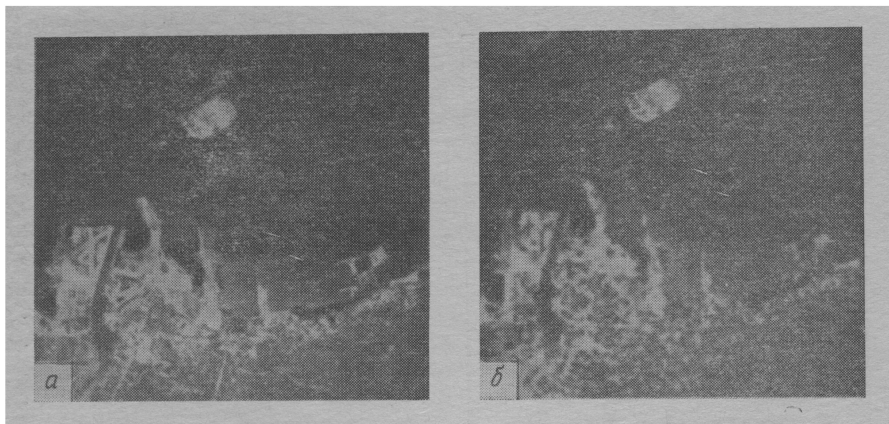


Рис. 21.4.1. Пример кодирования с переменной разрешающей способностью. Сокращение числа отсчетов на изображении, содержавшем первоначально 1024×1024 элементов, за счет снижения разрешающей способности в области фона в 64 раза; дальнейшее сжатие цифрового потока в восемь раз проведено путем кодирования с применением косинус-преобразования к блокам размером 256×256 элементов; общий коэффициент сжатия 44,5 : 1. Фотоснимки предоставлены У. Г. Ченом и Ч. Ридером (Ford Aerospace Corp., Пало-Альто, Калифорния).

а — оригинал; *б* — результат кодирования.

зрения. Снижение разрешающей способности на периферии изображения по сравнению с центральной зоной не отражается при этом на качестве воспроизведения основного объекта. К тому же и периферическому зрению человека, как известно, свойственно низкое пространственное разрешение по сравнению с фовеальной областью сетчатки. На рис. 21.4.2 показан пример дискретизации изображения при разделении поля зрения на четыре зоны с различным пространственным разрешением. Наиболее высокое разрешение в центральной зоне; в первом из окружающих ее поясов разрешение снижено вдвое, во втором поясе — вчетверо и т. д. Общее число элементов изображения при таком способе дискретизации составляет

$$N_R = 16R + 12\,(R/2) + 12\,(R/4) + 12\,(R/8) = 26{,}5R, \qquad (21.4.1)$$

где R — число элементов в зоне с максимальным разрешением. Если бы всюду на изображении поддерживалось такое же высокое

разрешение, как в центральной зоне, то общее число элементов увеличилось бы до $N_F = (32 \times 32)\, R$. Таким образом, полоса частот оказывается сокращенной в $N_F/N_R = 38{,}6$ раз.

Известно, что при восприятии подвижных объектов пространственная разрешающая способность зрительной системы человека значительно снижается — уловить тонкие детали объекта удается лишь в тех случаях, когда его движение в достаточной мере регулярно. Изображение быстро движущегося объекта в телевизионных системах обычно смазано, поскольку как камера, так и дисплей обладают определенной временно́й инерционностью [22]. Все это приводит к заключению, что при передаче в реальном масштабе времени можно без ущерба для качества изображения применить к движущимся относительно кадровой рамки объектам более грубую дискретизацию, сохраняя высокое разрешение лишь для статичных участков изображения [23—26]. К сожалению, такой подход значительно легче сформулировать, чем осуществить на практике. Основная проблема состоит в том, чтобы обнаружить сам факт движения объекта, а также указать числовые характеристики движения. В принципе пространственное разрешение для того или иного объекта должно быть обратно пропорционально его скорости.

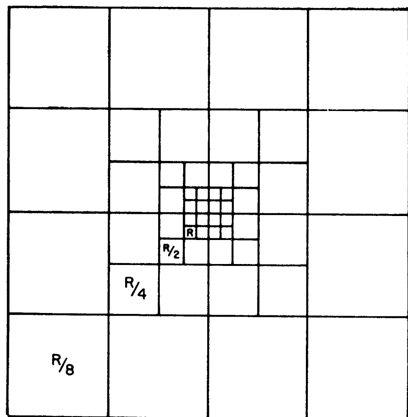


Рис. 21.4.2. Пример поля зрения с переменной пространственной разрешающей способностью.

Как показали эксперименты, проведенные Сейлером и Будрикисом [27], при телевизионной передаче в реальном масштабе времени наблюдателю требуется сравнительно длительный промежуток времени (700—800 мс) для зрительной адаптации к изображению при каждой смене сцены. В такие периоды времени возможно значительное снижение общей разрешающей способности телевизионной системы без ухудшения качества передачи.

21.4.2. УМЕНЬШЕНИЕ ЧАСТОТЫ КАДРОВ

Впечатление непрерывного движения может сохраняться при значительном уменьшении частоты кадров относительно величины 30 Гц, установленной телевизионным стандартом США. Однако при том способе передачи, который принят в телевидении, сниже-

ние частоты кадров приводит к мельканию изображения. Для устранения эффекта мелькания можно сохранить в малокадровой телевизионной системе высокую частоту полей, прибегая к многократному воспроизведению каждого кадра [28, 29]. Приемник такой системы должен быть, конечно, снабжен теми или иными средствами для запоминания кадров.

При наличии средств для запоминания кадров предел для снижения частоты передаваемых кадров связан с возникновением эффекта разрыва изображения при передаче сцен с движущимися объектами. Вопрос о том, можно ли с помощью специальных методов отодвинуть момент возникновения этого эффекта при снижении частоты кадров, был подвергнут экспериментальному изучению Болдуином [28]. Для экспериментов использовалась киносъемка со скоростью 24 кадра в секунду. Путем удаления определенных кадров исходного фильма изготавливался промежуточный малокадровый фильм. Затем посредством повторения сохраненных кадров изготавливался окончательный фильм, предназначенный для демонстрации с той же скоростью: 24 кадра в секунду. Последняя операция в простейшем варианте заключалась в том, что каждый сохраненный кадр печатался на пленке дважды или, если требовалось, несколько раз. Производились также пробы с более сложным комбинированием кадров. Для некоторых сцен, как, например, крупноплановое изображение разговаривающего человека, 12 кадров в секунду дают удовлетворительный результат. Однако этого оказалось недостаточно для сцены, показывающей движущийся в плоскости экрана поезд — в этом случае движение выглядело прерывистым. Дальнейшее снижение числа сохраненных кадров разрушало иллюзию непрерывности движения при демонстрации всех отснятых в этих экспериментах сюжетов. Что же касается усложненных методов комбинирования кадров, то они дали лишь незначительное улучшение по сравнению с простым повторением кадров.

Обычные телевизионные стандарты отводят почти 18 % периода строчной развертки и около 6 % периода кадровой развертки на обратный ход луча [15, стр. 2—22]. Уменьшение этих интервалов могло бы служить для сжатия спектра, хотя и ценой некоторого ухудшения качества изображения. В телевизионных системах с низкой скоростью развертки и передачи изображений таким способом можно получить небольшой выигрыш в полосе частот.

21.5. МЕТОДЫ ВРЕМЕННОГО ПЕРЕМЕЖЕНИЯ

Чересстрочная развертка, применяемая в обычном телевидении в качестве простого средства для сокращения полосы частот в отношении 2 : 1, была выбрана по тем соображениям, что она

не создает мельканий и эффекта разрыва изображения. Однако по отношению к телевизионным системам, приемники которых снабжены устройствами для запоминания кадров, вопрос о мельканиях из рассмотрения исключается. В таких системах единственным фактором, с которым приходится считаться при сокращении полосы частот методами временнóго перемежения, остается эффект разрыва изображения.

21.5.1. ПЕРЕДАЧА С ПЕРЕМЕЖЕНИЕМ СТРОК

Двукратное сокращение полосы частот достигается при чересстрочной развертке, когда в одном поле изображение развертывается по нечетным строкам, а в следующем поле — по четным. Частота полей в системе с чересстрочной разверткой соответствует частоте кадров в системе с последовательной разверткой строк. Развивая метод перемежения строк с коэффициентом 2 : 1, можно использовать перемежения с большими коэффициентами, передавая, например, каждую четвертую или восьмую строку в пределах одного поля. При рассмотрении систем с перемежением строк требуют особого внимания эффекты разрыва изображения и специфические иллюзии, мешающие восприятию изображения. При отсутствии перемежения телевизионное изображение движущегося объекта смазано и перемещается скачками. С введением перемежения эти особенности изображения сохраняются в том случае, если приемник системы не имеет специальных устройств для запоминания кадров или если частота кадров настолько мала, что по сравнению с длительностью передачи одного поля инерцией зрительного восприятия можно пренебречь. При этом, однако, сцена будет воспроизводиться с пониженным разрешением, как статическим, так и динамическим, поскольку всякий раз лишь некоторая часть общего числа строк будет предъявляться наблюдателю в один прием.

В системе с запоминанием полей на приемной стороне, когда предшествующие поля повторно воспроизводятся до момента их замены вновь поступающими полями, перемежение теоретически не снижает разрешающей способности. Однако на практике затруднительно сохранить с достаточной точностью равномерность взаимного расположения строк, относящихся к различным полям. Тенденция к спариванию строк при воспроизведении изображения существенно уменьшает фактическую разрешающую способность телевизионной системы.

В системе с перемежением движение поперек строк воспроизводится более мелкими скачками, чем в отсутствие перемежения; величина скачков уменьшается в том же отношении, какое принято в качестве коэффициента перемежения. Однако как следствие перемежения при этом происходит разрыв изображения. В ре-

зультате действия запоминающих устройств число строк, захватываемых изображением движущегося объекта, в определенной степени увеличивается. Этот эффект поясняется схемой на рис. 21.5.1. В отсутствие перемежения при передаче очередного поля каждая строка поля заменяется новой. В случае перемежения с коэффициентом 2 : 1 каждая строка поля лишь повторно

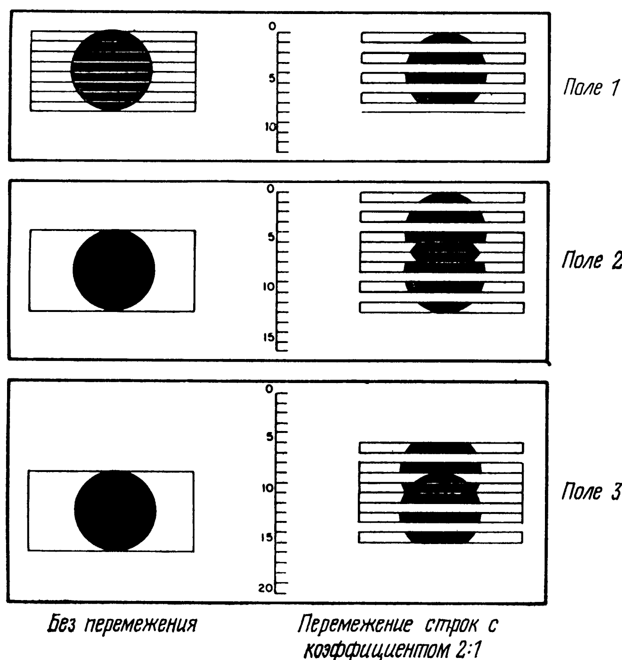


Рис. 21.5.1. Пространственный смаз изображения за счет перемежения строк.

воспроизводится при передаче очередного поля и заменяется новой при передаче последующего поля. Таким образом, в этом случае половина общего числа строк представляет «устаревшие» сведения. Как показано на рис. 21.5.1, в отсутствие перемежения движение круглого объекта сверху вниз передается от одного поля к другому отдельными скачками. Тот же объект, воспроизводимый системой с перемежением, оказывается размазанным в некоторой пространственной области, поскольку часть строк передает движение объекта с запозданием на одно поле. Такой же эффект возникает в системах с более высоким коэффициентом перемежения.

Другой эффект, связанный с перемежением строк, — это иллюзия сползания строк, возникающая в случае последовательного воспроизведения полей. На рис. 21.5.2 для объяснения этого

эффекта показано расположение строк, принадлежащих последовательным полям, при значениях коэффициента перемежения 2 : 1 и 3 : 1. Цифрами 1, 2, 3 обозначены строки соответствующих полей. Пусть перемежение производится с коэффициентом 3 : 1, как показано на рис. 21.5.2, *б*. После того как наблюдателю было предъявлено поле 3, появляется поле 1. Строки поля 1 оказываются одним интервалом ниже и двумя интервалами выше строк поля 3. В результате возникает зрительное впечатление, что строки поля 3 просто несколько сдвинулись в направлении сверху вниз, образовав строки поля 1. В свою очередь строки последующего поля 2 выглядят как результат небольшого сдвига в том же направлении строк поля 1 и т. д. Направление сползания не зависит от того, в какой последовательности передаются строки поля — сверху вниз или снизу вверх. Стандартная чересстрочная развертка (коэффициент перемежения 2 : 1, рис. 21.5.2, *а*) исключает появление такой иллюзии, поскольку строки полей 1 и 2 отделены друг от друга равными интервалами. Вполне возможно, что эффект сползания может быть ослаблен путем строчной развертки с псевдослучайной последовательностью передачи строк.

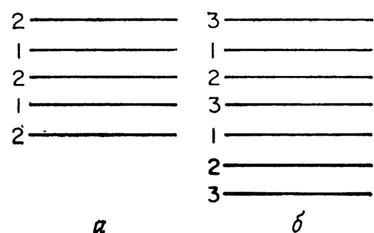


Рис. 21.5.2. Строки развертки, принадлежащие различным полям, для случаев перемежения с коэффициентами 2 : 1 (*а*) и 3 : 1 (*б*).

21.5.2. ПЕРЕДАЧА С ПЕРЕМЕЖЕНИЕМ ЭЛЕМЕНТОВ

Основной вариант перемежения с коэффициентом 2 : 1 может быть применен не только к строкам (как в вещательном телевидении), но и к произвольно разграниченным по вертикали и горизонтали фрагментам изображения [30, 31]. С самой общей точки зрения перемежение состоит в том, что кадр разбивается на несколько последовательно передаваемых полей, каждое из которых содержит определенный набор элементов изображения. Обычно поля формируются так, чтобы они не перекрывали друг друга. Сокращение полосы частот обеспечивается тем, что на передачу каждого поля отводится такой же промежуток времени, какой обычная телевизионная система затрачивает на кадр. Если, например, кадр разбивается на 16 неперекрывающихся полей, то эквивалентная частота кадров уменьшится в 16 раз и коэффициент сжатия спектра составит 16 : 1. По сообщению Дейча [30], система с коэффициентом перемежения 16 : 1, действующая к тому же при малой частоте кадров и пониженном разрешении, способна

вести передачу по каналу с полосой частот всего лишь 10 кГц.

Рис. 21.5.3 дает представление о рисунках перемежения для ряда систем с перемежением элементов. Для сравнения приве-

Без перемежения
Скорость развертки R
Сокр. полосы частот 1:1
Кадр содержит 1 поле

Рисунок поля

A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A

Рисунок кадра в целом

A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A

Перемежение строк 2:1
Скорость развертки R/2
Сокр. полосы частот 2:1
Кадр содержит 2 поля

A	A	A	A	A	A	A	A
A	A	A	A	A	A	A	A

A	A	A	A	A	A	A	A
B	B	B	B	B	B	B	B
A	A	A	A	A	A	A	A

Вертикальное перемежение 2:1
Скорость развертки R
Сокр. полосы частот 2:1
Кадр содержит 2 поля

A		A		A		A	
A		A		A		A	
A		A		A		A	

A	B	A	B	A	B	A	B
A	B	A	B	A	B	A	B
A	B	A	B	A	B	A	B

Перемежение элементов 2:1
Скорость развертки R
Сокр. полосы частот 2:1
Кадр содержит 2 поля

A		A		A		A	
	A		A		A		A
A		A		A		A	

A	B	A	B	A	B	A	B
B	A	B	A	B	A	B	A
A	B	A	B	A	B	A	B

Перемежение элементов 2:1 и
Перемежение строк 2:1
Скорость развертки R/2
Сокр. полосы частот 4:1
Кадр содержит 4 поля

A		A		A		A	
A		A		A		A	

A	C	A	C	A	C	A	C
D	B	D	B	D	B	D	B
A	C	A	C	A	C	A	C

Перемежение элементов 4:1
Скорость развертки R
Сокр. полосы частот 4:1
Кадр содержит 4 поля

A				A			
		A				A	
	A				A		
			A				A
A				A			

A	C	B	D	A	C	B	D
B	D	A	C	B	D	A	C
D	A	C	B	D	A	C	B
C	B	D	A	C	B	D	A
A	C	B	D	A	C	B	D

Перемежение элементов 8:1
Скорость развертки R
Сокр. полосы частот 8:1
Кадр содержит 8 полей

A							
			A				
						A	
	A						
				A			
							A
		A					
					A		

A	E	B	F	C	G	D	H
G	D	H	A	E	B	F	C
B	F	C	G	D	H	A	E
H	A	E	B	F	C	G	D
C	G	D	H	A	E	B	F
E	B	F	C	G	D	H	A
D	H	A	E	B	F	C	G
F	C	G	D	H	A	E	B

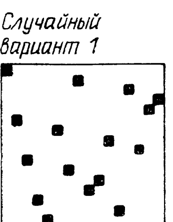
Рис. 21.5.3. Способы перемежения при передаче изображений.

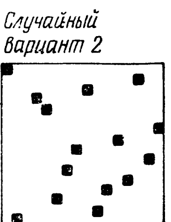
дены также рисунки поля в телевизионных системах без перемежения, с перемежением строк и с вертикальным перемежением. Следует отметить, что перемежение элементов, как правило, не стоит сочетать с перемежением строк в связи с проблемами разрыва изображения и сползания строк, о которых говорилось выше.

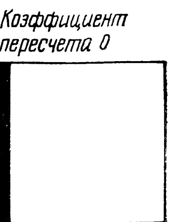
Для каждой из систем, представленных на рис. 21.5.3, сокращение полосы частот относительно системы без перемежения указано для фиксированной скорости развертки, измеряемой числом R пройденных в единицу времени элементов кадра. Перемежение элементов с коэффициентом 2 : 1 дает единственные возможные варианты рисунка первого поля и рисунка следующего поля. Положение меняется при перемежении более высокого порядка. Коэффициенты перемежения 4 : 1, 8 : 1, 16 : 1 и т. д. предоставляют большой выбор вариантов при составлении рисунка отдельных полей и последовательности их воспроизведения. От того, как произведен этот выбор, зависит степень проявления специфических визуальных эффектов, связанных с перемежением.

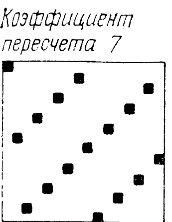
При перемежении элементов наблюдаются оптические иллюзии двух типов — наложение узора и скольжение элементов. Наложение узора проявляется в виде ложной текстуры изображения, зависящей от рисунка перемежения. Узор появляется в тех случаях, когда вместе с элементами очередного поля воспроизводятся повторно не все элементы других полей. Иллюзия сползания или скольжения объясняется стробоскопическими эффектами при последовательном воспроизведении полей; элементы изображения как бы бегут при этом по экрану. Продуманный рисунок полей и удачное их последовательное расположение могут свести проявление обоих эффектов к минимуму. С точки зрения практического удобства желательно, чтобы отдельные поля до некоторой степени повторяли друг друга по своей структуре, а последовательность полей также можно было периодически повторять в процессе передачи. Обычно при этих условиях еще большее внимание следует уделять эффектам наложения узора и скольжения элементов.

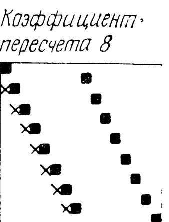
Рассмотрим в качестве примера кодирующую систему с перемежением элементов, где каждый кадр разделяется на 16 неперекрывающихся полей. Рисунок перемежения повторяется в каждом из блоков размером 16×16 элементов. На рис. 21.5.4 в пределах одного блока показан ряд расположений элементов поля для системы с коэффициентом перемежения 16 : 1. Первые два рисунка перемежения построены с помощью таблицы случайных чисел при тех, однако, условиях, что поле включает первый элемент каждого блока, и любой столбец и любая строка поля в пределах блока содержат ровно один элемент. Далее приведена серия базовых рисунков перемежения, формирование которых сводится к простым операциям счета импульсов в процессе передачи. Счет строк и счет столбцов ведется следующим образом: единичное приращение номера строки вызывают, например, в зависимости от выбранного коэффициента пересчета каждый второй или каждый третий импульс, поступивший в счетчик столбцов. Содержимое счетчика строк определяет вертикальную координату элемента поля; при единичном увеличении этой координаты содержимое

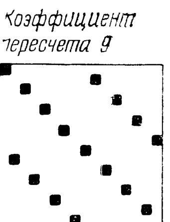


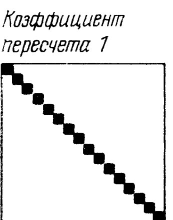


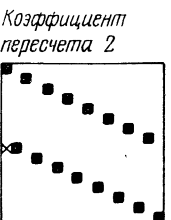




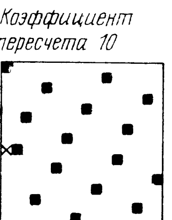




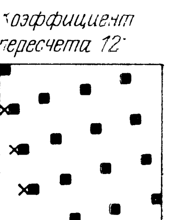


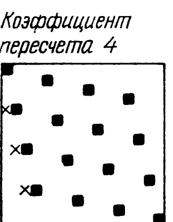
















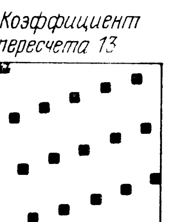






Рис. 21.5.4. Варианты рисунка поля в пределах блока размером 16×16 при перемежении элементов.

счетчика столбцов поступает в регистр горизонтальной координаты элемента. После достижения последней строки или последнего столбца блока в результате переполнения соответствующего счетчика происходит его сброс, и счет начинается заново. В том случае, когда очередной элемент поля оказывается в столбце или в строке, где уже имеется другой элемент, содержимое счетчика столбцов увеличивается дополнительно на единицу. Такие случаи отмечены на рис. 21.5.4 знаком ×. Наиболее благоприятное впечатление оставляет узор, получаемый при выборе для коэффициента пересчета значения 6. Элементы поля располагаются при этом равномерно и в меньшей степени вызывают ощущение регулярной линейчатой структуры.

Каждый базовый рисунок можно многими способами дополнить неперекрывающимися последующими полями. Наиболее простой подход состоит, однако, в том, чтобы получать дальнейшие поля в результате применения к базовому рисунку операций сдвига. С этой точки зрения любой узор на рис. 21.5.4 можно рассматривать как таблицу для формирования последовательных полей: данному номеру поля ставится в соответствие некоторый сдвиг по вертикали. Можно одновременно варьировать сдвиг по горизонтали, определяющий положение начального элемента базового рисунка. Для формирования последовательности полей счетчики строк и столбцов могут включать в общий счет также импульсы полей с тем или иным коэффициентом пересчета. Понятно, что малые значения коэффициентов пересчета дадут не самый лучший результат. Кроме того, для максимального подавления эффекта наложения узора следует избегать одинаковых коэффициентов пересчета при формировании базового рисунка и последовательности воспроизведения полей. Если инерция зрения, послесвечение люминофора или память какого-либо накопителя данных не обеспечивают в достаточной мере запоминания полей, то при наблюдении неподвижных сцен разрешающая способность снижается в том же отношении, какое принято в качестве коэффициента перемежения. В малокадровых телевизионных системах введение устройства запоминания полей становится необходимостью, поскольку инерция зрения и послесвечение люминофора приемной трубки уже не могут помочь в повышении разрешения. При наличии в приемнике запоминающих устройств все принятые ранее поля, не обновленные к данному моменту вновь поступающими полями, могут многократно извлекаться из памяти и воспроизводиться на дисплее с достаточно высокой частотой. Таким образом, располагая памятью, вмещающей полный комплект полей, можно полностью сохранить разрешение при воспроизведении неподвижных сцен. Однако такая память приводит к нежелательному размазыванию движущихся объектов. Этот эффект проявляется тем сильнее, чем больше полей содержится

в запоминающем устройстве и, следовательно, чем выше разрешение для неподвижных сцен. Подбирая число запоминаемых полей, можно установить некоторый компромисс в отношении качества воспроизведения подвижных и неподвижных участков сцены. Другие компромиссные решения приходится принимать в связи с тем, что при увеличении объема памяти сильнее проявляется эффект наложения узора, в то время как эффект скольжения элементов ослабляется. Наконец, еще одну проблему при выборе параметров системы составляет снижение заметности хаотического фона, возникающего в том случае, когда при каждом акте воспроизведения полей на дисплее обновляются за счет памяти не все элементы изображения. Если на месте таких элементов воспроизводится уровень белого или уровень черного, то это воспринимается как дополнительная помеха на изображении. Подводя итог, следует заключить, что метод перемежения элементов может служить достаточно простым средством сокращения частот в тех практических случаях, когда подвижность передаваемых сцен относительно невелика и при воспроизведении движущихся объектов допускается определенная степень пространственного размазывания.

21.6. РАЗВЕРТКА С ПЕРЕМЕННОЙ СКОРОСТЬЮ

Метод сокращения полосы частот путем управления скоростью развертки состоит в следующем. Видеодатчик ускоряет процесс развертки на участках изображения с низкой детальностью и замедляет его на участках с высокой детальностью [33—37]. Основная часть интересующих наблюдателя сведений заключена в контурах изображения. При обычной телевизионной передаче вероятность появления контура на протяжении строки развертки достаточно мала; основная часть периода строки приходится в среднем на пологие участки рельефа яркости. Управляя скоростью развертки таким образом, чтобы увеличенная доля общего времени приходилась на замедленную передачу сведений о контурах, а соответственно уменьшенная доля времени — на ускоренную передачу сведений о пологих участках, можно сократить полосу частот видеосигнала. Такой видоизмененный видеосигнал сохраняет соответствие исходному изображению и при наличии на приемной стороне сведений о скорости развертки может быть после некоторой обработки преобразован в обычный видеосигнал. В принципе необходимые для восстановления изображения сведения о скорости развертки содержатся в самом же видоизмененном видеосигнале, так что отпадает необходимость в дополнительном канале для их передачи.

Блок-схема телевизионной системы с переменной скоростью развертки представлена на рис. 21.6.1. Видеодатчик включен

в контур обратной связи, модулирующий скорость развертки согласно соотношению

$$V = dx/dt = V_M - K|de/dt|, \qquad (21.6.1)$$

где x — расстояние вдоль строки развертки, e — напряжение сигнала яркости на выходе камеры, V_M — максимально допусти-

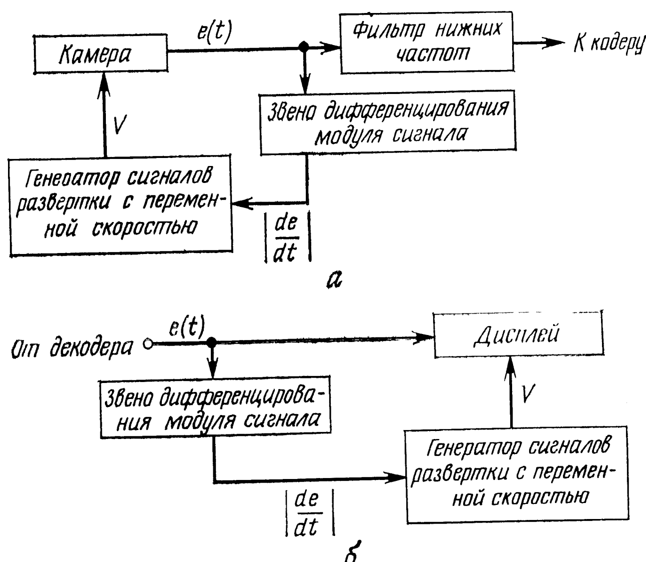


Рис. 21.6.1. Система с переменной скоростью развертки.
а — передатчик; *б* — приемник.

мая скорость развертки, K — постоянный коэффициент. Поскольку

$$|de/dt| = V|de/dx|, \qquad (21.6.2)$$

то

$$V = V_M/(1 + K|de/dx|). \qquad (21.6.3)$$

Пусть

$$\eta = V_M/V_N \qquad (21.6.4)$$

есть отношение максимальной скорости развертки к нормальной скорости, T — время, затрачиваемое в среднем на прохождение строки длиной L в процессе развертки с переменной скоростью, а T_N — соответствующее время при развертке с нормальной скоростью V_N. Отношение этих промежутков времени дается формулой

$$T/T_N = [1 + (\eta - 1) D_F]/\eta, \qquad (21.6.5)$$

где D_F есть показатель детальности изображения, определяемый как

$$D_F = (1/L) \int_0^L |de/dx| \, dx. \qquad (21.6.6)$$

Показатель детальности изображения в виде шахматного поля составляет 1, так что в этом случае ускорения развертки не происходит. Для изображения в виде белого поля, как и следовало ожидать, $D_F = 1/\eta$. Как показывают измерения, для изображений естественного происхождения величина D_F лежит в пределах от 3 до 5 % [35]. Отношение T/T_N есть показатель сокращения времени, затрачиваемого в среднем на передачу строки. Полоса частот может быть сокращена путем надлежащей дискретизации сигнала, получаемого в ходе развертки с переменной скоростью, и распределения отсчетов в интервале времени T_N. Если, например, $D_F = 0{,}05$ и $\eta = 10$, то полоса частот может быть сокращена в отношении 7 : 1.

Поскольку управление скоростью развертки основывается на операции дифференцирования, отношение сигнал/шум на выходе телевизионной камеры приходится поддерживать на уровне около 60 дБ [4] во избежание вариаций в длине воспроизводимых строк. В обычных телевизионных системах стандартная величина отношения сигнал/шум составляет ~30 дБ. Но кодирующим системам с переменной скоростью развертки свойствен и другой существенный недостаток. Для того чтобы увеличить выигрыш в полосе частот, ее верхняя граница при проектировании системы выбирается в расчете на передачу шахматного поля с максимальным размахом яркости, порождающего высокочастотный видеосигнал. Но шахматное поле с уменьшенным размахом яркости дает видеосигнал еще более высокой частоты. На передающей стороне системы производится фильтрация видеосигнала, что ухудшает воспроизведение мелких, малоконтрастных деталей изображения. Данное обстоятельство в сочетании с жесткими требованиями к величине отношения сигнал/шум ограничивает практическую применимость кодирующих систем с переменной скоростью развертки.

21.7. МУЛЬТИПЛЕКСНЫЕ МЕТОДЫ КОДИРОВАНИЯ

В мультиплексных системах телефонной связи, позволяющих с помощью нескольких телефонных каналов обслуживать множество переговоров, используется тот факт, что речевой сигнал содержит большое число пауз. Предоставляя имеющиеся каналы источникам сигнала, находящимся в данный момент в активной фазе, такие системы способны действовать с достаточно малой

вероятностью наложения разговоров. Важную роль при этом играют нерегулярный характер сигналов и отсутствие корреляции между различными источниками.

Методы мультиплексной передачи речевых сигналов были перенесены в область передачи изображений [39—41]. В простейшей из систем такого типа, предназначенной для передачи черно-белых факсимильных изображений, производится синхронная строчная развертка нескольких различных документов. Если в данный момент времени только на одном из этих документов встретился черный элемент, по каналу связи передается двоичная комбинация, указывающая номер соответствующего документа. Если же все документы одновременно дают сигнал белого (фон), то передается набор нулей. Следовательно, в процессе одновременной развертки N документов при переходе на позицию очередного элемента передается одно из $N + 1$ кодовых слов. Равномерный код, длина слова которого составляет $\log_2 (N + 1)$ двоичных разрядов, при надлежащем выборе числа N ($N = 7$, 15, 31, ...) сокращает объем передаваемых данных в

$$C = N/[\log_2 (N + 1)] \tag{21.7.1}$$

раз. Если $N = 7$, то сокращение выразится отношением 2,3 : 1. Разумеется, система допускает ошибки в передаче, когда несколько документов одновременно дают сигнал черного. На каждом из таких документов, за исключением одного, при воспроизведении соответствующий черный элемент будет заменен белым элементом. Как показывает цифровое моделирование [39], при одновременной передаче семи машинописных страниц, заполненных черными элементами в среднем на 10 %, около 35 % всех черных элементов оказываются стертыми. Тем не менее даже при такой сравнительно высокой вероятности стирания лишь около 4 % знаков не удается правильно распознать без привлечения контекста.

Хуанг и Третьяк [42] предложили подойти к мультиплексной передаче изображений иначе. Для каждого из N документов устанавливается свой псевдослучайный закон развертки. Все N сигналов поступают в логическую схему ИЛИ, на выходе которой формируется «единица», если хотя бы один из сигналов имеет уровень черного, а в противном случае — «нуль». Переданная последовательность нулей и единиц одновременно поступает на все N приемных устройств. Каждое из них повторяет псевдослучайную последовательность развертки соответствующего передающего устройства. При поступлении очередной единицы все приемные устройства воспроизводят черный элемент, однако только одно из них при этом наверняка не сделает ошибки. Поскольку ошибочно принятые черные элементы имеют псевдослучайное пространственное распределение, они создают на изобра-

жении помеху типа снега. Переданные таким способом документы достаточно легко поддаются прочтению при одновременной передаче четырех изображений [42]. Получаемые изображения могут быть затем значительно улучшены путем фильтрации, устраняющей обособленные черные элементы.

ЛИТЕРАТУРА

1. Huang T. S., PCM Picture Transmission, *IEEE Spectrum*, **2**, 12, 57—60 (December 1965).
2. Special Issue on Redundancy Reduction, *Proc. IEEE*, **55**, 3 (March 1967). [Имеется перевод: ТИИЭР, тематический вып. «Сокращение избыточности», 1967, т. 55, № 3.]
3. Huang T. S., Schreiber W. F., Tretiak O. J., Image Processing, *Proc. IEEE*, **59**, 11, 1586—1609 (November 1971).
4. Special Issue on Signal Processing for Digital Communications, *IEEE Trans. Commun. Tech.*, COM-**19**, 6, Part I (December 1971).
5. Special Issue on Digital Picture Processing, *Proc. IEEE*, **60**, 7 (July 1972). [Имеется перевод: ТИИЭР, тематический вып. «Обработка изображений при помощи цифровых вычислительных машин», 1972, т. 60, № 7.]
6. Huang T. S., Tretiak O. J., Eds., Picture Bandwidth Compression, Gordon and Breach, New York, 1972.
7. Habibi A., Robinson G. S., A Survey of Digital Picture Coding, *IEEE Computer*, **7**, 5, 21—34 (May 1974).
8. Wilson J. C., Television Engineering, Pitman, London, 1937.
9. Ives H. E., Johnsrud A. L., Television in Colors by a Beam Scanning Method, *J. Opt, Soc. Am.*, **20**, 11—21 (January 1930).
10. Goldmark P. C. et al., Color Television — Part I., *Proc. IRE*, **30**, 4, 162—182 (April 1942).
11. Goldmark P. C., et al., Color Television — Part II, *Proc. IRE*, **31**, 9, 465—578 (September 1943).
12. McLean F. C., Worldwide Color Television Standards, *IEEE Spectrum*, **3**, 6, 59—60 (June 1966).
13. Glassford G. M., Fundamentals of Television Engineering, McGraw-Hill, New York, 1955.
14. McIlwain K., Dean C. E., Eds., Principles of Color Television, Wiley, New York, 1956.
15. Fink D. G., Ed., Television Engineering Handbook, McGraw-Hill, New York, 1957.
16. McIlwain K., Requisite Color Bandwidth for Simultaneous Color-Television Systems, *Proc. IRE*, **40**, 909—912 (August 1952).
17. Bailey W. F., Hirsch C. J., Quadrature Cross Talk in NTSC Color Television, *Proc. IRE*, **42**, 84—90 (January 1954).
18. de France H., The Sequential-Simultaneous Colour Television System, *Acta Electronics*, **2**, 392—397 (1957—1958).
19. Townsend B., PAL Colour Television, Cambridge University Press, Cambridge, England, 1970.
20. Howson E. A., Bell D. A., Reduction of Television Bandwidth by Frequency Interlace, *J. Brit. IRE*, **20**, 2, 127—136 (February 1960).
21. Maier J. L., Gardenhire L., Redundant Area Coding System (REARCS), Proceedings of the International Telemetering Conference, Los Angeles, California, Vol. 8, October 1972, pp. 301—314.
22. Haskell B. G., Mounts F. W., Candy J. C., Interframe Coding of Videotelephone Pictures, *Proc. IEEE*, **60**, 7, 792—800 (July 1972). [Имеется перевод: Хэскилл, Маунтс, Кэнди. Межкадровое кодирование видеотелефонного изоб-

ражения. — ТИИЭР, тематический вып. «Обработка изображений при помощи цифровых вычислительных машин», 1972, т. 60, № 7, с. 47—58.]
23. Seyler A. J., The Coding of Visual Signals to Reduce Channel-Capacity Requirements, *Proc. Inst. Elec. Eng.*, **109**, Part C, 676—684 (September 1962).
24. Pease R. F. W., Limb J. O., Exchange of Spatial and Temporal Resolution in Television Coding, *Bell Syst. Tech. J.*, **50**, 1, 191—200 (January 1971).
25. Limb J. O., Pease R. F. W., A Simple Interframe Coder for Video Telephony, *Bell. Syst. Tech. J.*, **50**, 6, 1877—1888 (July—August 1971).
26. Candy J. C. et al., Transmitting Television as Clusters of Frame-to-Frame Differences, *Bell Syst. Tech. J.*, **50**, 6, 1889—1917 (July—August 1971).
27. Seyler A. J., Budrikis Z. L., Detail Perception After Scene Changes in Television Image Presentations, *IEEE Trans. Inf. Theory*, IT-**11**, 1, 31—43 (January 1965).
28. Baldwin M. W., Demonstration of Some Visual Effects of Using Frame Storage in Television Transmission, IRE Convention Record, 1958, p. 107.
29. Brainard R. C., Mounts F. W., Prasada B., Low Resolution TV: Subjective Effects of Frame Repetition and Picture Replenishment, *Bell Syst. Tech. J.*, **46**, 1, 261—271 (January 1967).
30. Deutsch S., Pseudo-Random Dot Scan Television Systems, *IEEE Trans. Broadcasting*, BC-**11**, 1, 11—21 (July 1965).
31. Gouriet G. G., Dot Interlaced Television, *Electronic Eng.*, **24**, 290, 166—171 (April 1952).
32. Haantjes J., Teer K., Multiplex Television Transmission: Subcarrier and Dot Interlace Systems, *Wireless Eng.*, Part I, **31**, 215—233 (September 1954); Part II, **31**, 266—273 (October 1954).
33. Wrught E. E., Velocity Modulation in Television, *Proc. Phys. Soc. (London)*, **46**, 512—514 (July 1934).
34. Bedford L. H., Puckle O. S., A Velocity Modulation Television System, *J, IEE*, **75**, 63 (1934).
35. Cherry E. C., Gouriet G. G., Some Possibilities for the Compression of Television Signals by Recoding, *Proc. IEE*, **100**, Part II, 3—18 (1953).
36. Beddoes M. P., Experiments with Slope Feedback Coder for TV Compression, *IRE Trans. Broadcasting*, **BC-7**, 2, 12—28 (March 1961).
37. Beddoes M. P., Two Channel Method for Compressing Bandwidth of Television Signals, *Proc. IEE*, **110**, 2, 369—374 (February 1963).
38. Prince M. D., Honnel M. A., TV Image Reproduction by Use of Velocity-Modulation Principles, *Proc. IRE*, **39**, 3, 265—268 (March 1951).
39. Seitzer D., Class F., Stucki P., An Experimental Approach to Video Bandwidth Compression by Multiplexing, *IEEE Trans. Commun. Tech.*, **COM-17**, 5, 564—568 (October 1969).
40. Stucki P., Limits of Instantaneous Priority Multiplexing Applied to Black-and-White Pictures, *IEEE Trans. Commun. Tech.*, **COM-19**, 2, 169—177 (April 1971).
41. Class F., Video Bandwidth Compression by Multiplexing: Experimental Results, *IEEE Trans. Commun. Tech.*, **COM-19**, 3, 371—372 (June 1971).
42. Huang T. S., Tretiak O. J., A Pseudorandom Multiplex System for Facsimile Transmission, *IEEE Trans. Commun. Tech.*, **COM-16**, 3, 436—438 (June 1968).

Глава 22

ЦИФРОВОЕ КОДИРОВАНИЕ ИЗОБРАЖЕНИЙ С ПОЭЛЕМЕНТНОЙ ОБРАБОТКОЙ

В этой главе рассматриваются цифровые методы кодирования изображений с поэлементной обработкой. Описываются системы кодирования с импульсно-кодовой модуляцией, статистические методы кодирования, метод пополнения кадров и кодирование с предсказанием. Рассматриваются также вопросы проектирования кодирующих систем, основанных на линейных методах предсказания.

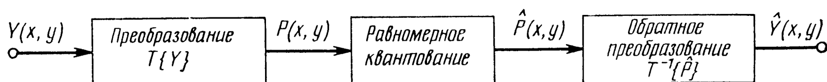
22.1. СИСТЕМЫ КОДИРОВАНИЯ С ИМПУЛЬСНО-КОДОВОЙ МОДУЛЯЦИЕЙ

В обычной системе кодирования изображений с импульсно-кодовой модуляцией (ИКМ) непрерывный видеосигнал преобразуется в последовательность квантованных отсчетов. Эти отсчеты представляются затем наборами нулей и единиц одинаковой длины, называемыми кодовыми словами. Очевидно, что расход двоичных цифр при передаче можно снижать путем уменьшения общего числа уровней квантования. Когда качество изображения, создаваемого такой системой, оценивается с помощью некоторого количественного критерия, число уровней квантования определяется тем минимумом, при котором обеспечивается заданное качество. Если же качество изображения оценивается субъективно, то число уровней квантования приходится сохранять достаточно большим во избежание появления ложных контуров. Этот эффект, возникающий вследствие скачкообразных изменений яркости от одного уровня квантования к другому, наиболее заметен в тех частях восстановленного изображения, где яркость оригинала меняется медленно. Чтобы предотвратить появление ложных контуров на одноцветных изображениях, обычно требуется 50 и более уровней квантования. В соответствии с этим в системах кодирования одноцветных изображений с применением ИКМ число уровней квантования находится обычно в пределах от 64 до 256, так что каждому элементу изображения соответствует 6—8-разрядное двоичное кодовое слово [1—4]. Для кодирования цветных изображений в обычных системах ИКМ отводится по 6—8 двоичных единиц на передачу каждой из трех координат цвета — красного, зеленого и синего. Разработан также ряд сравнительно простых ме-

тодов цифрового кодирования изображений, позволяющих значительно снизить по сравнению с обычными системами ИКМ число уровней квантования без существенного увеличения заметности ложных контуров.

22.1.1. КВАНТОВАНИЕ СО СЖАТИЕМ ДИАПАЗОНА ЯРКОСТЕЙ

Вопрос об оптимальной шкале квантования для случаев, когда объектом квантования является единственный передаваемый па-



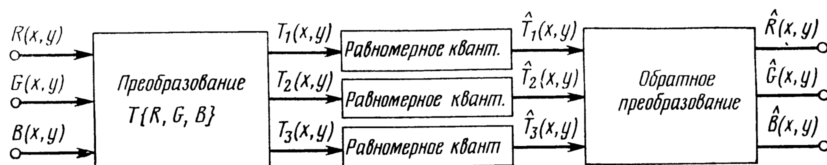


Рис. 22.1.1. Квантование изображений со сжатием диапазона яркостей.
а — для одноцветных изображений; *б* — для цветных изображений.

раметр или совокупность параметров, рассмотрен в шестой главе. Было определено расположение пороговых уровней и уровней квантования, сводящее к минимуму среднеквадратическую ошибку квантования. Как выяснилось, пороговые уровни и уровни квантования в диапазоне возможных значений передаваемой функции должны быть, вообще говоря, расположены неравномерно. Квантование со сжатием, описанное в шестой главе, представляет собой по существу способ реализации неравномерного квантования. Как показано на рис. 22.1.1, в процессе квантования со сжатием входная функция подвергается нелинейному преобразованию, затем равномерному квантованию и, наконец, обратному нелинейному преобразованию.

Пусть требуется провести квантование в расчете на минимальную заметность ложных контуров при визуальном восприятии одноцветного изображения. В этом случае пороговые уровни и уровни квантования должны быть расположены так, чтобы любой шаг квантования соответствовал одинаковому приращению светлоты — величины, определенной в разд. 7.3, которая связывает воспринимаемую яркость с физической яркостью. Другими словами, нужно равномерно квантовать шкалу светлоты, выполнив

поэлементное преобразование яркости в светлоту [4]. Такая процедура квантования может быть распространена и на случай цветного изображения, если воспользоваться геодезической метрикой как функцией, определяющей нелинейное преобразование.

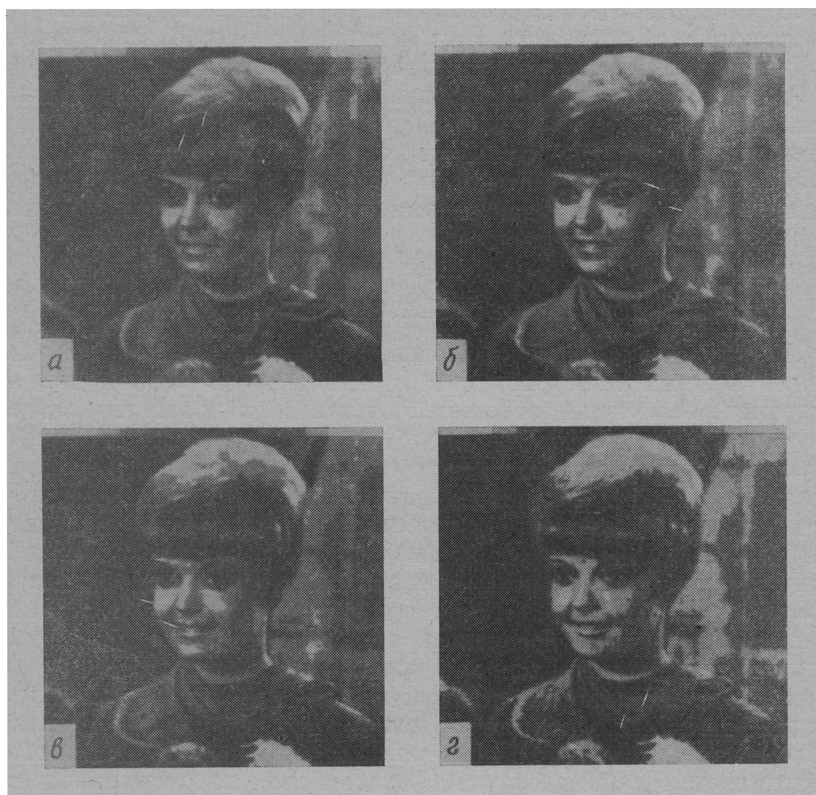


Рис. 22.1.2. Примеры квантования на 8 уровней с применением поэлементных преобразований сжатия.

а — квантование без сжатия; *б* — логарифмическое сжатие; *в* — сжатие по закону квадратного корня; *г* — сжатие по закону кубического корня.

Другой подход к квантованию изображения состоит в том, чтобы перед операцией равномерного квантования отобразить яркость изображения или тройку координат цвета в пространство сигналов зрительной системы человека [5]. При этом нелинейное преобразование непосредственно определяется моделями зрительной системы человека, описанными во второй главе.

Рис. 22.1.2 иллюстрирует применение различных поэлементных преобразований сжатия для квантования одноцветных изображений. Эти результаты получены путем цифрового моделирования:

яркость элементов исходного изображения подвергалась равномерному 12-разрядному квантованию в прецизионном устройстве ввода в ЭВМ. После нелинейного поэлементного преобразования заново производилось равномерное квантование на 8 уровней и таким образом моделировалось квантование со сжатием. Эти эксперименты показывают, что нелинейное преобразование сжатия до некоторой степени снижает заметность ложных контуров и различные законы преобразования приводят к приблизительно равноценным результатам. По сообщению Фрэя [6], квантование цветных изображений со сжатием, основанным на модели цветного зрения, позволяет без потери качества изображения вдвое уменьшить число уровней квантования по сравнению с равномерным квантованием без сжатия.

22.1.2. ПСЕВДОШУМОВОЕ КВАНТОВАНИЕ

Визуальный эффект квантования, проявляющийся в образовании ложных контуров, может быть в значительной степени устранен, если перед операцией квантования добавить к видеосигналу малую псевдослучайную составляющую (псевдошум) и одновременно вычесть ее из сигнала, получаемого на выходе квантователя [7—12]. Эта процедура, называемая в честь ее автора [7] модуляцией Робертса, позволяет квантовать видеосигнал, затрачивая всего лишь 3 двоичные единицы на каждый отсчет и не допуская при этом появления сколько-нибудь заметных ложных контуров. Вместе с тем такой процесс квантования характеризуется повышенной среднеквадратической ошибкой и искажениями, воспринимаемыми как «снег» на изображении.

Блок-схема системы псевдошумового квантования показана на рис. 22.1.3. Входной отсчет, величина x которого представляет яркость элемента изображения или некоторую функцию яркости, объединяется с псевдослучайным сигналом n_T, и полученная сумма z поступает на квантователь, имеющий равномерную шкалу с шагом Δ. Возможные значения псевдослучайного сигнала n_T распределены равномерно в интервале от $-\Delta/2$ до $\Delta/2$ с плотностью

$$p(n_T) = 1/\Delta. \qquad (22.1.1)$$

Сигнал n_R, поступающий от другого источника псевдошума с равномерной плотностью распределения, вычитается из выходного сигнала квантователя w. В системе псевдошумового квантования сигналы n_T и n_R формируются датчиками случайных чисел, работа которых может быть синхронизирована для получения идентичных псевдослучайных последовательностей.

Основную идею метода иллюстрирует рис. 22.1.4. Предположим, например, что некоторому отрезку строки изображения



Рис. 22.1.3. Схема процесса псевдошумового квантования.

Рис. 22.1.4. Пример псевдошумового квантования.

а — обычное квантование; *б* — псевдошумовой сигнал; *в* — сумма входного и шумового сигналов; *г* — квантованная сумма входного и шумового сигналов; *д* — результат псевдошумового квантования

соответствует линейно нарастающий сигнал (рис. 22.1.4, *а*). Любой переход через пороговый уровень вызывает при воспроизведении этого сигнала скачок от одного уровня квантования к соседнему. Как раз к перепадам такого характера глаз проявляет особую чувствительность, и именно они ответственны за возникновение ложных контуров. Но положение меняется, когда к исходному сигналу добавляется псевдошумовая составляющая с равномерным распределением (рис. 22.1.4, *б*) и лишь затем производится квантование (рис. 22.1.4, *г*). Псевдослучайная составляющая вносит в результат квантования дополнительные колебания вокруг уровня квантования, которым был бы представлен восстановленный сигнал в отсутствие псевдошума. Как видно, общий размах скачков видеосигнала после квантования возрастает. Далее, как видно из рис. 22.1.4, *д*, псевдошумовая компонента, первоначально добавленная к входному сигналу, вычитается из результата квантования. Эта операция ведет к определенному сближению восстановленного сигнала с тем усредненным сигналом, который существовал бы в отсутствие шумового воздействия на процесс квантования. Как можно видеть из приведенного примера, резкие скачки уровня уступают в данном случае место, вообще говоря, уменьшенным перепадам, сопровождающим каждый переход от элемента к элементу.

Следуя Робертсу, действие системы в целом можно оценить среднеквадратическим отклонением выходного сигнала от входного, т. е. величиной

$$\mathscr{E} = \int_0^1 \int_0^1 (x - y)^2 p(x, y)\, dx\, dy, \tag{22.1.2}$$

полагая при этом, что в результате нормировки значения входных отсчетов заключены в интервале от 0 до 1 и что $p(x, y)$ есть совместная плотность вероятности для переменных на входе и выходе. Мерой дополнительного шума, обнаруживаемого на изображении в результате псевдошумового воздействия на процесс квантования, может служить усредненная по всем возможным значениям входного сигнала дисперсия выходного сигнала относительно его условного математического ожидания. Дисперсия этой дополнительной ошибки квантования определяется как

$$V = \int_0^1 \int_0^1 (y - \bar{y}_x)^2 p(x, y)\, dx\, dy, \tag{22.1.3}$$

где

$$\bar{y}_x = \int_0^1 y p(y \mid x)\, dy \tag{22.1.4}$$

есть условное среднее выходного сигнала. Наконец, отклонение условного среднего выходного сигнала относительно входного сигнала определяется как

$$D = \int_0^1 \int_0^1 (x - \bar{y}_x)^2 p(x, y)\, dx\, dy. \qquad (22.1.5)$$

В отсутствие псевдошума среднеквадратическая ошибка квантования принимает минимальное значение, а дисперсия дополнительной ошибки обращается в нуль. При введении малых шумовых компонент n_T и n_R как среднеквадратическая ошибка, так и дисперсия V дополнительного шума на выходе системы несколько возрастут, однако отклонение D уменьшится. Получаемые при этом изображения производят значительно лучшее впечатление, что объясняется способностью глаза к пространственному усреднению малых колебаний яркости. Данные, характеризующие систему псевдошумового квантования в случае изображений с равномерным распределением яркости, приведены в табл. 22.1.1. Все эти данные нормированы относительно среднеквадратической ошибки $\mathscr{E}$ в отсутствие псевдошума (т. е. при обычной ИКМ).

Таблица 22.1.1

Расчетные характеристики системы псевдошумового квантования для изображений с равномерным распределением яркости. (Размах псевдошумового сигнала равен шагу квантования, b — число уровней квантования.)

	Нормированная среднеквадратическая ошибка	Нормированная дисперсия шума квантования	Уклонение
ИКМ	1	0	1
Добавление шумовой компоненты после квантования	2	1	1
Добавление шумовой компоненты перед квантованием	$2 - 2^{-b}$	$2(1 - 2^{-b})$	2^{-b}
Добавление шумовой компоненты перед квантованием и ее вычитание после квантования	$1 + 2^{-b}$	1	2^{-b}

22.1.3. КВАНТОВАНИЕ С УЛУЧШЕННОЙ ПЕРЕДАЧЕЙ ГРАДАЦИЙ ЯРКОСТИ

В 1966 г. Бизиньяни, Ричардс и Уилан [13] предложили новый метод квантования изображений, названный квантованием с улучшенной передачей градаций яркости. Для борьбы с ложными контурами метод предусматривает малое возмущение яркости при

квантовании очередного элемента, находящееся в детерминированной зависимости от яркости предшествующего по ходу развертки элемента. В случае шестиразрядного квантования «искусственное» кодовое слово образуется суммированием по модулю 8 обычного кодового слова данного элемента с содержимым трех младших разрядов предшествующего кодового слова. На операцию суммирования налагается запрет в том случае, когда кодовое слово данного элемента в трех старших разрядах содержит только единицы. Наконец, при передаче изображения искусственное кодовое слово сокращают до трех старших разрядов, так что рас-

Таблица 22.1.2

Пример кодирования в системе с улучшенной передачей градаций яркости при затрате 3 дв. ед./эл. [13]

Порядковый номер элемента	Уровень элемента	Код элемента	Искусственный код	Сокращенный код	Восстанавливаемый уровень	Стандартная трехразрядная ИКМ
0[a]	—		000 000	—		
1	12	001 100	001 100	001	8	8
2	12	001 100	010 000	010	16	8
3	13	001 101	001 101	001	8	8
4	13	001 101	010 010	010	16	8
5	10	001 010	001 100	001	8	8
6	13	001 101	010 001	010	16	8
7	9	001 001	001 010	001	8	8
8	9	001 001	001 011	001	8	8
9	15	001 111	010 010	010	16	8
10	13	001 101	001 111	001	8	8
11	19	010 011	011 010	011	24	16
12	38	100 110	101 000	101	40	32
13	40	101 000	101 000	101	40	40
14	24	011 000	011 000	011	24	24
15	10	001 010	001 010	001	8	8
16	10	001 010	001 100	001	8	8
17	10	001 010	001 110	001	8	8
18	10	001 010	010 000	010	16	8
19	10	001 010	001 010	001	8	8
20	10	001 010	001 100	001	8	8
21	10	001 010	001 110	001	8	8

[a] Начальное условие

ход двоичных цифр снижается вдвое. Пример кодирования в соответствии с описанным алгоритмом дан в табл. 22.1.2. Последний столбец показывает, как были бы представлены элементы изображения в случае передачи обычным трехразрядным кодом. Система с улучшенной передачей градаций яркости затрачивает всего лишь

три двоичные единицы на элемент, но более частые флуктуации уровня при воспроизведении изображения создают эффект размывания ложных контуров в результате пространственного усреднения при зрительном восприятии. Восстановление изображения сопровождается некоторым увеличением среднеквадратической ошибки по сравнению с обычной ИКМ. Изображения, переданные системой с улучшенной передачей градаций яркости, в значительной степени освобождаются от ложных контуров, но обнаруживают небольшую зернистость.

22.1.4. СИСТЕМА С ГРУБОЙ-ТОНКОЙ ШКАЛОЙ КВАНТОВАНИЯ

В системе с грубой-тонкой шкалой квантования, также разработанной Бизиньяни, Ричардсом и Уиланом [13, 14], предусматривается равномерное квантование с использованием 64 уровней и формирование для каждого элемента изображения предварительного шестиразрядного двоичного кодового слова. Три старших разряда этого слова содержат грубую оценку уровня элемента, а три младших разряда — уточняющую информацию. Если изменение уровня при передаче изображения сказывается на грубой оценке, то ею и ограничиваются как окончательной величиной уровня, и содержимое трех соответствующих разрядов составляет передаваемый код. С другой стороны, если грубая оценка сохраняется без изменения, то передается лишь уточняющая трехразрядная часть слова с тем, чтобы, комбинируя ее с грубой трехразрядной оценкой для предшествующего элемента, восстановить на приемной стороне шестиразрядный код для воспроизведения уровня очередного элемента. Следует подчеркнуть, что на передачу каждого элемента затрачиваются лишь 3 двоичные единицы. На практике вопрос о выборе между грубой и тонкой шкалами квантования решается несколько сложнее, чем описано выше. Решение для очередного элемента принимается с помощью специально выработанного набора правил на основе данных об уровнях трех предшествующих элементов [13, 14]. В высшей степени адаптивный характер работы делает систему с грубой-тонкой шкалой квантования достаточно трудным объектом для математического анализа. Что же касается субъективных испытаний, то квантование с грубой-тонкой шкалой дает изображения такого же качества, что и шестиразрядная ИКМ. Некоторое ухудшение отмечается лишь вблизи резких перепадов яркости, где эффективность подавления ложных контуров снижается.

Изображения, показанные на рис. 22.1.5, позволяют сравнить четыре различных метода: трехразрядную ИКМ, псевдошумовое квантование, квантование с улучшенной передачей полутонов и квантование с грубой-тонкой шкалой. Ложные контуры, харак-

терные для ИКМ, на других трех изображениях не обнаруживаются. Тем не менее три последние системы вносят свои искажения яркости. В случае псевдошумового квантования эти искажения выглядят как аддитивный шум, в двух других случаях они приобретают более локализованный характер.

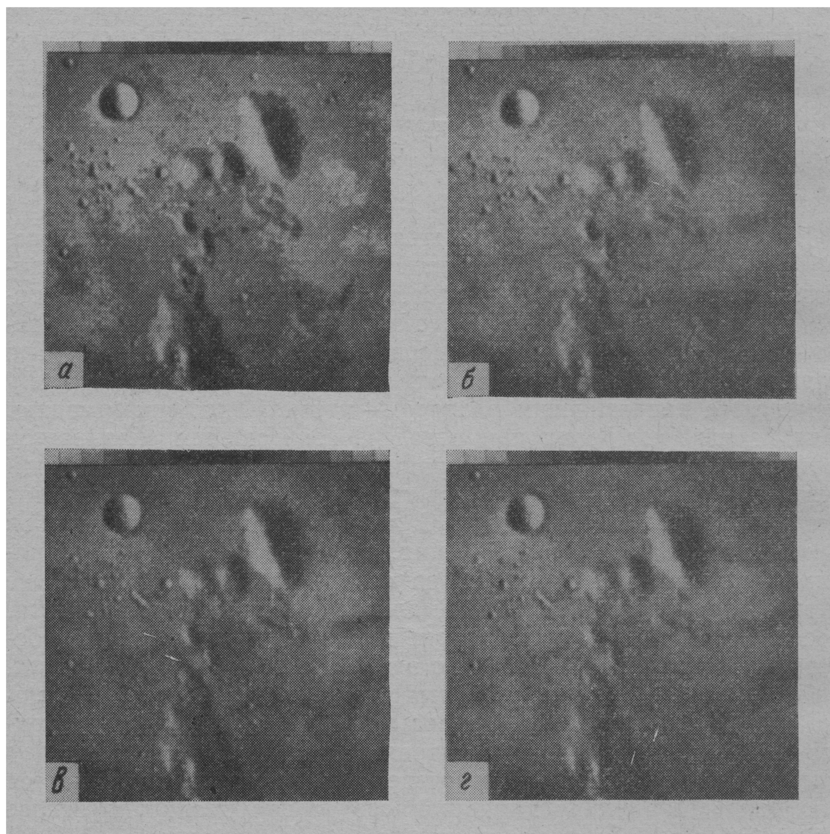


Рис. 22.1.5. Применение экономных методов квантования для кодирования с затратой 3 дв. ед./эл.

а — ИКМ; *б* — псевдошумовое квантование; *в* — квантование с улучшенной передачей градаций яркости; *г* — квантование с грубой-тонкой шкалой.

22.1.5. КВАНТОВАНИЕ С ЧАСТОТНЫМ РАЗДЕЛЕНИЕМ

В систему квантования с частотным разделением [15] входят электрические фильтры, посредством которых сигнал, получаемый в результате телевизионной развертки, расщепляется на две компоненты — низкочастотную и высокочастотную. Квантование низкочастотной компоненты производится в расчете на передачу

пяти-шестиразрядным двоичным кодом, а высокочастотной — трехразрядным кодом. Такое различие основывается на предположении, что глаз менее чувствителен к искажениям яркости на тех участках изображения, которым соответствуют быстрые изменения видеосигнала. Система весьма проста в практическом осуществлении, но получаемая с ее помощью экономия в затрате двоичных цифр невелика, в то время как субъективное качество передаваемых изображений в некоторой степени снижается.

22.1.6. УМЕНЬШЕНИЕ ОШИБОК КВАНТОВАНИЯ В СИСТЕМАХ ИКМ

Действие системы кодирования с ИКМ основано на последовательной независимой передаче элементов изображения. Однако тот факт, что между элементами имеется пространственная корреляция, дает возможность уменьшить ошибку квантования методом векторного квантования, кратко описанным в гл. 6. При этом для совокупности J элементов, образующих вектор $\mathbf{f}$, требуется указать подходящий набор многомерных ячеек квантования D_j, разграниченных пороговыми поверхностями, а для каждой такой ячейки — вектор уровней квантования $\mathbf{r}_j$. Определение оптимальной системы ячеек квантования представляет собой, вообще говоря, чрезвычайно сложную вычислительную задачу. Поэтому на практике ограничиваются прямоугольными ячейками. Если предположить, что система ячеек квантования задана, то оптимальный вектор квантования определяется как

$$\mathbf{r}_j = \int_{D_j} \mathbf{f} p(\mathbf{f})\, d\mathbf{f} \Bigg/ \int_{D_j} p(\mathbf{f})\, d\mathbf{f}, \tag{22.1.6}$$

где $p(\mathbf{f})$ есть совместная плотность вероятности элементов вектора $\mathbf{f}$. С помощью итерационной процедуры Хунс [16] решил эту задачу, определяя уровень квантования элемента по значениям пороговых уровней четырех соседних элементов (левого, правого, верхнего и нижнего). Решение получено в предположении, что совместная плотность вероятности пяти элементов, участвующих в процессе восстановления, хорошо описывается совместной плотностью распределения Рэлея коррелированных случайных величин. Рис. 22.1.6 позволяет сравнить результаты, получаемые в обычной системе ИКМ, с изображениями, восстановленными методом векторного квантования. По сравнению с обычной ИКМ ложные контуры во всех случаях проявляются значительно слабее, а среднеквадратическая ошибка уменьшается. Существует также возможность подавления ложных контуров путем пространственного сглаживания восстанавливаемого изображения в результате свертки с некоторой сглаживающей функ-

цией или фильтрации в спектральной области. Однако такие операции снижают возможности системы в воспроизведении мелких деталей изображения.



Рис. 22.1.6. Уменьшение ошибок квантования в системе ИКМ методом обобщенного квантования [16].
Кодирование с затратой 1 дв. ед./эл.:
а — обычная ИКМ; *б* — восстановление методом векторного квантования. Кодирование с затратой 2 дв ед./эл; *в* — обычная ИКМ; *г* — восстановление методом векторного квантования.

22.2. СТАТИСТИЧЕСКИЕ МЕТОДЫ КОДИРОВАНИЯ

Как показывают статистические данные о распределении яркости элементов изображений и соответствующие оценки энтропии, изображения естественных сцен обладают значительной избыточностью. Для целей данного рассмотрения избыточность можно определить как общее число двоичных единиц, затрачиваемых при обычной ИКМ, за вычетом энтропии изображения в целом. Рассмотрим в качестве примера одноцветное изображение, состоящее из $N \times N$ элементов. Если яркость каждого элемента

квантована на B уровней, то общее число различаемых изображений составляет

$$T = B^{N^2}. \tag{22.2.1}$$

Однако в реальных условиях далеко не любое из этих изображений будет иметь сколько-нибудь значительную вероятность появления. Пусть p_i есть вероятность появления i-го из возможных изображений. Энтропия ансамбля изображений, определяющая минимальную среднюю затрату двоичных цифр на изображение, составит

$$H = -\sum_{i=1}^{T} p_i \log_2 p_i. \tag{22.2.2}$$

Для кодирования изображений в системах ИКМ количество уровней квантования

$$B = 2^b \tag{22.2.3}$$

выбирается обычно так, чтобы b было целым числом (от 6 до 8). В результате система ИКМ затрачивает в общей сложности

$$N_B = bN^2 \tag{22.2.4}$$

двоичных единиц для того, чтобы каждому из N^2 элементов сопоставить кодовое слово стандартной длины, состоящее из b двоичных разрядов. По своей величине энтропия H естественных изображений на много порядков меньше числа N_B двоичных цифр, затрачиваемых системой ИКМ. Отсюда следует, разумеется, что ИКМ как средство кодирования изображений обладает чрезвычайно низкой эффективностью.

Возникает вопрос, существует ли в таком случае вообще метод кодирования, с помощью которого можно свести затрату двоичных цифр до минимума, определяемого энтропией. Теория отвечает на этот вопрос утвердительно, однако практическое осуществление такой кодирующей системы в настоящее время не представляется возможным.

Рассмотрим следующую схему прямого решения задачи оптимального кодирования. Пусть изображению с большой вероятностью появления p_i ставится в соответствие относительно короткое кодовое слово, а длинные кодовые слова сопоставляются изображениям с более низкой вероятностью появления. Существует процедура оптимального кодирования, приводящая к кодам (известным как коды Хаффмена [17]), средняя длина слова которых весьма близко подходит к нижней границе, определяемой нормированной величиной энтропии H/N^2. Предположим, что все слова такого кода могут быть собраны для хранения в кодовой книге. В таком случае, когда предъявляется конкретное изображение, процедура кодирования сводится просто к извлечению

соответствующего кодового слова из полного перечня изображений. Декодер такой системы должен произвести обратную операцию — найти переданное кодовое слово в книге и отобрать изображение, подлежащее воспроизведению. И хотя такая непосредственная процедура кодирования оптимальна, ее осуществление связано с чрезвычайно большим объемом операций. Пусть, например, изображение содержит 256×256 элементов, а яркость каждого элемента квантована на 256 уровней. Тогда общее число возможных изображений равно $10^{158\ 000}$! Понятно, что просмотр кодовой книги такого объема находится за пределами практических возможностей. Можно, конечно, попытаться разделить изображение на малые блоки размером, скажем, 4×4 элементов. Однако даже в этом случае число различных блоков приближается к 10^{40}, что также не оставляет надежд на практически выполнимую процедуру поиска в кодовой книге. Но имеется и другое существенное препятствие для кодирования с помощью подобной процедуры: чтобы составить подходящий код, необходимо знать вероятности p_i появления каждого из возможных изображений. Практическое измерение или оценка необходимых априорных вероятностей представляет собой очень сложную задачу. Отсутствие надежных вероятностных данных обычно резко снижает эффективность работы кодирующей системы.

Хотя перспективы оптимального статистического кодирования выглядят достаточно мрачно, с помощью субоптимальных методов удается все-таки исключить по крайней мере некоторую часть избыточности изображений, достигая при этом значительного снижения требований к пропускной способности канала связи. В последующих разделах дается описание нескольких методов статистического кодирования изображений. Некоторые из этих методов полностью исключают потери информации, т. е. не вносят искажений в передаваемые изображения. Другие же допускают малый находящийся под контролем уровень искажений в обмен на более эффективное сжатие диапазона передаваемых частот. Ряд описываемых ниже методов статистического кодирования может служить основой для построения подсистем других систем кодирования изображений.

22.2.1. ПОЭЛЕМЕНТНОЕ СТАТИСТИЧЕСКОЕ КОДИРОВАНИЕ

Сущность простейшего метода статистического кодирования изображений заключается в том, что каждому элементу в отдельности на основании его уровня квантования ставится в соответствие определенное кодовое слово из кодовой книги. Кодирование будет эффективным, если уровням с большой вероятностью появления будут соответствовать короткие кодовые слова, и наоборот, элементы с реже встречающимися уровнями будут переда-

ваться более длинными кодовыми словами. Если должным образом организовать этот процесс, средняя длина передаваемых кодовых слов сравняется с энтропией элемента изображения.

В качестве первого шага при составлении кода необходимо произвести моделирование, оценку или измерение вероятностей появления элементов с различными уровнями для изображений определенного типа. Запишем вероятность появления n-го уровня квантования r_n в виде

$$P(n) = \Pr\{F(j, k) = r_n\}. \tag{22.2.5}$$

В процессе кодирования каждому уровню квантования ставится в соответствие некоторое кодовое слово, состоящее из $b(n)$ двоичных разрядов, так что средняя длина кодовой комбинации составляет

$$L_c = \sum_{n=0}^{L-1} P(n)\, b(n). \tag{22.2.6}$$

Эффективность E_f кода представляет собой отношение энтропии $H\{F(j, k)\}$ элемента (вычисленной в предположении взаимной независимости элементов) к средней длине кодового слова. Таким образом,

$$E_f = H\{F(j, k)\}/L_c. \tag{22.2.7}$$

Существует ряд высокоэффективных кодов, подходящих для поэлементного статистического кодирования изображений. В приложении 3 дается описание двух таких кодов: Шеннона—Фано и Хаффмэна. Последний всегда реализует наивысшую достижимую для данного источника сообщений эффективность. Эти коды, вообще говоря, характеризуются переменной длиной, и поэтому их применение требует буферного накопителя данных для выравнивания скорости передачи. Но статистическое кодирование связано и с другой проблемой общего характера. При формировании кода основываются на определенных предположениях о распределении вероятностей, характеризующем источник сообщений. В том случае, когда эти предположения не выполняются, в работе кодирующей системы может наступить резкое ухудшение. И действительно, плохо согласованный с источником код может привести даже к увеличению средней затраты двоичных цифр по сравнению с системой ИКМ, оперирующей кодом фиксированной длины.

Для одноцветных изображений естественного происхождения, квантованных на 64 уровня, энтропия элемента, рассматриваемого независимо от других элементов, изменяется в пределах от 4 до 6 дв. ед./эл. Такой относительно малый выигрыш по сравнению с шестиразрядной ИКМ в большинстве случаев не оправдывает усложнений, возникающих при использовании кодов с переменной длиной.

22.2.2. КОДИРОВАНИЕ С УЧЕТОМ ПРЕДШЕСТВУЮЩЕГО ЭЛЕМЕНТА

Для большей части естественных изображений характерна высокая степень корреляции между соседними элементами вдоль строки развертки, т. е. существует большая избыточность. Для одноцветных изображений, квантованных на 64 уровня, энтропия значения предшествующего элемента составляет около 2—3 дв. ед./эл. Если исходить из этой величины, то система ИКМ, затрачивающая на передачу каждого элемента 6 дв. ед., обладает существенной избыточностью. Статистический кодер, способный устранить эту избыточность, в принципе достаточно прост. Каждому из L уровней квантования, представляющему яркость предшествующего элемента, соответствуют условное распределение вероятностей $P(j|k)$ и связанный с этим распределением набор кодовых слов. Кодирующая система обращается к k-й (уровень квантования предшествующего элемента) кодовой книге и выбирает из нее j-е (уровень квантования данного элемента) кодовое слово. При этом придется либо хранить в памяти L^2 кодовых слов, либо предусмотреть алгоритм формирования кодового слова для каждой пары (j, k). Если производить квантование на 64 уровня, то кодовая книга будет содержать в общей сложности 4096 позиций. На практике во многих случаях кодирующая система такой сложности оказывается неприемлемой. Переходя далее от пар элементов к тройкам, можно, как правило, снизить удельный расход двоичных цифр дополнительно на 0,5 дв. ед./эл., однако кодовая книга должна будет содержать уже L^3 позиций. Следует к тому же отметить, что ошибка в кодовом слове, возникшая при передаче по каналу связи, искажает при воспроизведении изображения не только соответствующий элемент, но и все последующие элементы в пределах строки. Такое распространение искажений может быть ограничено путем периодической передачи на протяжении строки кодовых слов, определяющих уровень элемента независимо от соседних элементов.

Таблица 22.2.1

Пример безошибочного кодирования разницы в уровнях между соседними элементами

Разность уровней D	Кодовое слово
0	1
+1	0100
−1	0101
+2	0110
−2	0111
+3	00100
−3	00101
+4	00110
−4	00111
$\|D\| \geq 5$	000 + 6-разр. кодовое слово

Один из вариантов кодирования с учетом предшествующего элемента состоит в кодировании разности уровней соседних элементов; уровень первого элемента строки кодируется при этом независимо [18—20]. Если имеется L уровней квантования, то разность уровней может принимать одно из $2L-1$ значений. Поскольку вероятность появления разностей с большой абсолютной величиной относительно мала, можно существенно упростить кодер, не ухудшая значительно качество его работы. За малыми разностями закрепляются индивидуальные кодовые слова; если же разность превышает установленный порог, то формируется кодовое слово, соответствующее уровню очередного элемента, причем этот код объединяется с ключевой кодовой комбинацией, выделяющей его среди закодированных разностей. Табл. 22.2.1 иллюстрирует применение для этой цели кода Шеннона—Фано. Эффективность кодов такого типа может достигать 90 % при относительной простоте реализации, что обусловливает их практическую ценность [18].

Таблица 22.2.2

Пример кодирования разницы в уровнях элементов, допускающего погрешность в один шаг квантования

Разность уровней D	*Кодовое слово*
0 ± 1	1
3 ± 1	0100
6 ± 1	0101
-3 ± 1	0110
-6 ± 1	0111
9 ± 1	00100
12 ± 1	00101
-9 ± 1	00110
-12 ± 1	00111
15 ± 1	0001000
18 ± 1	0001001
21 ± 1	0001010
24 ± 1	0001011
-15 ± 1	0001100
-18 ± 1	0001101
-21 ± 1	0001110
-24 ± 1	0001111
$\|D\|\geq26$	0000 + 8-*разр. кодовое слово*

Все рассмотренные до сих пор статистические коды обеспечивают передачу без потерь информации. Однако к системе статистического кодирования не обязательно предъявлять требование полного сохранения информации. Табл. 22.2.2 иллюстрирует метод кодирования разности уровней соседних элементов, не делающий различий в кодовых комбинациях для разностей, величина которых варьирует в пределах плюс—минус одного шага квантования.

Как уже отмечалось, одна из основных проблем статистического кодирования заключается в том, что эффективность кодирования сохраняется лишь при достаточно хорошем согласовании кода с распределением вероятностей, характеризующим источник сообщений. В связи с этим Райс и Плаунт [21] разработали адаптивную систему статистического кодирования изображений, преду-

сматривающую текущий контроль реального распределения частот в процессе кодирования и выбор среди некоторого набора статистических кодов варианта, наилучшим образом соответствующего источнику сообщений.

22.3. КОДИРОВАНИЕ СЕРИЙ

Кодирование серий [22—30] — сравнительно простой метод передачи, основанный на сравнении уровней соседних элементов вдоль строки развертки. Серией называют последовательность элементов, уровни которых мало отличаются друг от друга. Соседние серии отделяются большим скачком уровня — контуром. При таком способе кодирования передается либо некоторая функция яркости последнего элемента серии, либо функция значения разностного сигнала на границе серий; кроме того, указывается позиция последнего элемента серии. Если эта позиция определяется числом элементов от начала строки, то говорят о кодировании концов серий. Позицию последнего элемента данной серии можно также определить относительно последнего элемента предыдущей серии. Такой метод называют кодированием длин серий.

В цифровых системах передачи сведения об уровне и о расположении серий представляются в виде кодовых комбинаций. Передача этих сведений по мере появления окончаний серий в процессе непрерывной развертки изображения дает возможность снизить энергию сигнала, однако ширина полосы частот при этом не уменьшается. Сжатие спектра возможно лишь при накоплении данных о сериях и передаче их со скоростью, соответствующей средней частоте появления серий. Эта цель может быть достигнута с помощью запоминающего блока в считывающей системе либо путем управления скоростью развертки с таким расчетом, чтобы серии следовали с заданным интервалом времени [31].

Кодирование концов серий имеет тот недостаток, что для указания позиции на строке, составленной из нескольких сотен элементов, требуется хотя и фиксированная по длине, но многоразрядная кодовая комбинация. Что же касается кодирования длин серий, то средняя длина кодовых комбинаций, определяющих положение границ серий, здесь сравнительно невелика. Тем не менее переменная длина этих комбинаций затрудняет кодирование. В связи с этим кодирование позиций может быть видоизменено следующим образом: для соответствующих кодовых комбинаций назначается укороченная стандартная длина, а тем самым и максимально возможная длина серии. Если окончание серии оказывается за пределами этого максимального интервала, то ее остаток передается как «псевдосерия». Псевдосерия воспринимается как серия с нулевым приращением яркости относительно предыдущей серии. Этот метод в свою очередь имеет недостаток,

который состоит в том, что общее число серий при описании изображения увеличивается, поскольку для многих серий число разрядов в кодовой комбинаций, указывающей длину серии, оказывается чрезмерным и приходится дополнительно передавать псевдосерии по мере их появления. Однако надлежащий выбор длины максимально допустимого интервала, соответствующий статистическим свойствам сообщений, может уменьшить избыточность такого рода.

22.3.1. КОДИРОВАНИЕ ДЛИН СЕРИЙ; ТЕОРЕТИЧЕСКОЕ РАССМОТРЕНИЕ [32—36]

Простая, но достаточно точная модель процесса кодирования длин серий получается в предположении, что обнаружение перепада яркости на данном элементе изображения есть независимое событие. Пусть вероятность обнаружения перепада p одинакова для всех элементов, а Z обозначает длину серии, т. е. число составляющих ее элементов ($Z = 1$ соответствует двум перепадам, следующим подряд). Распределение вероятностей длин серий представляет собой геометрическое распределение

$$P\{Z = j\} = q^{j-1}p, \tag{22.3.1}$$

где $q = 1 - p$. Длина серии составляет в среднем

$$E\{Z\} = 1/p. \tag{22.3.2}$$

Энтропия распределения вероятностей длин серий определяется как

$$H\{Z\} = -\sum_{j=1}^{\infty} q^{j-1} p \log_2 \{q^{j-1}p\} = -[q \log_2 q + p \log_2 p]/p. \tag{22.3.3}$$

Предположим теперь, что максимально возможная длина серий ограничена величиной M элементов. При этом распределение вероятностей длин серий примет следующий вид:

$$P\{Z_M = j\} = q^{j-1}p + q^M \delta(j - M), \quad j = 1, 2, \ldots, M. \tag{22.3.4}$$

Средняя длина серий составит теперь

$$E\{Z_M\} = \sum_{j=1}^{M} jq^{j-1}p + Mq^M = (1/p) + Mq^M, \tag{22.3.5}$$

а энтропия распределения вероятностей длин серий будет равна

$$H\{Z_M\} = -(1/p)\left(1 - q^{M-1}\right)(p \log_2 p + q \log_2 q). \tag{22.3.6}$$

На рис. 22.3.1 показана зависимость средней длины серий от вероятности появления контура. Вообще говоря, для каждого значения p можно указать такое значение максимальной длины серий, которое минимизирует энтропию распределения вероят-

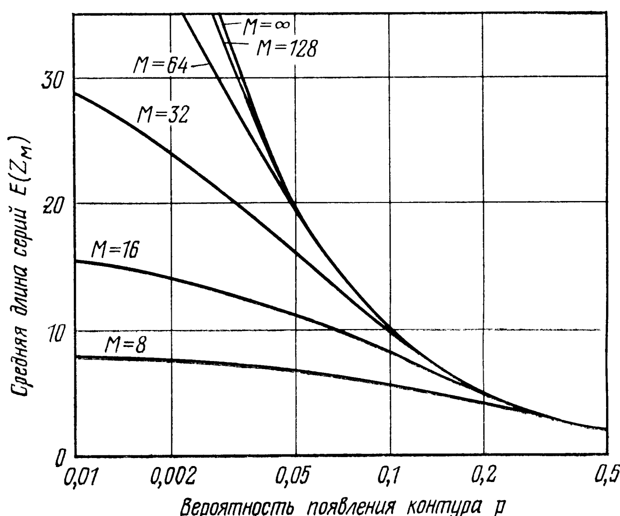


Рис. 22.3.1. Средняя длина серий. (M — предел, установленный для длины серий.)

ностей длин серий. Однако главный интерес представляет не сама энтропия, а ее доля, приходящаяся на элемент изображения:

$$H\{E\} = H\{Z_M\}/E\{Z_M\}. \qquad (22.3.7)$$

На рис. 22.3.2 приведены графики зависимости этой относительной величины энтропии от вероятности появления контура при $M = 8$, 16, а также при $M = \infty$. Графики со всей очевидностью демонстрируют, сколь незначителен проигрыш, связанный с ростом относительной величины энтропии за счет ограничения длины серий. На рис. 22.3.2 представлена также графически величина

$$L_c = G/E\{Z_M\}, \qquad (22.3.8)$$

выражающая среднюю затрату двоичных единиц на элемент при передаче длин серий равномерным кодом с помощью G-разрядных кодовых комбинаций; $M = 2^G$ (дв. ед.). Как видно, в случаях, представляющих наибольший практический интерес ($0,1 \leqslant p \leqslant 0,2$), графики величины L_c вплотную подходят к графикам величины $H\{E\}$.

Рассмотрим теперь кодирование сведений о перепадах уровня на контурах. В простейшем случае, когда объектом кодирования

является уровень элемента, следующего за любым значительным перепадом яркости, суммарная энтропия распределений вероятностей подлежащих кодированию величин составляет

$$H\{E+A\} = [H\{F(j,\ k)\} + H\{Z_M\}]/E\{Z_M\}. \qquad (22.3.9)$$

Эта величина, очевидно, зависит от числа уровней квантования яркости элементов изображения, поскольку число уровней непо-

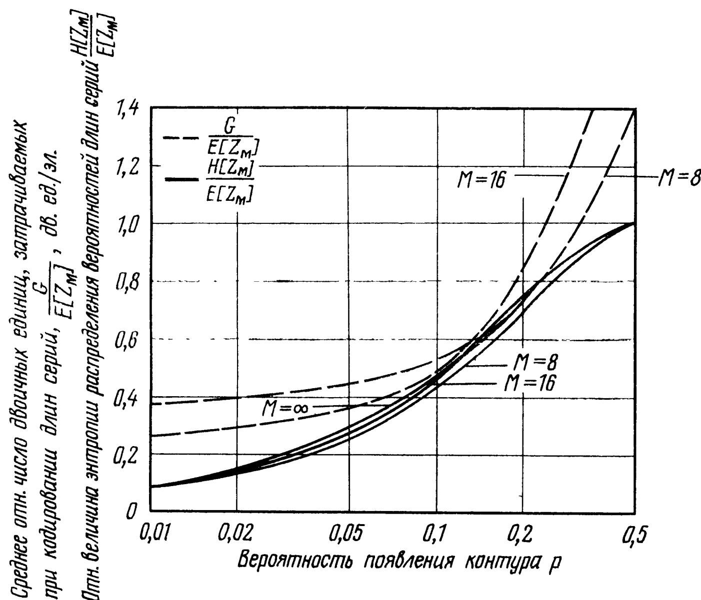


Рис. 22.3.2. Эффективность кодирования длин серий для случая двухградационных изображений.

средственно влияет на слагаемое $H\{F(j,\ k)\}$. Однако даже в большей степени, хотя и не прямым образом, а через посредство вероятности появления контура, это число воздействует на энтропию распределения длин серий и на среднюю длину серий. Когда число уровней превышает 16, вероятность появления перепада становится столь значительной, что процесс кодирования теряет всякий практический смысл.

Представляет интерес выяснение эффективности системы, основанной на кодировании длин серий, в случае передачи как длин, так и перепадов яркости удобными на практике кодами с фиксированной длиной слова. Если на задание перепада яркости и длины серии отводятся соответственно A и G двоичных единиц, то средняя относительная длина кодового слова в расчете на элемент изображения составит

$$L_c = (A+G)/E\{Z_M\}. \qquad (22.3.10)$$

Графики этой величины для случая, когда $A = 3$, и предельная длина серии $M = 2^G$ приведены на рис. 22.3.3. Огибающая кривых указывает минимум, до которого снижается средний расход двоичных цифр на элемент изображения, когда предельная длина серии выбирается оптимальной для данного диапазона вероятностей появления контура. Если эта вероятность составляет

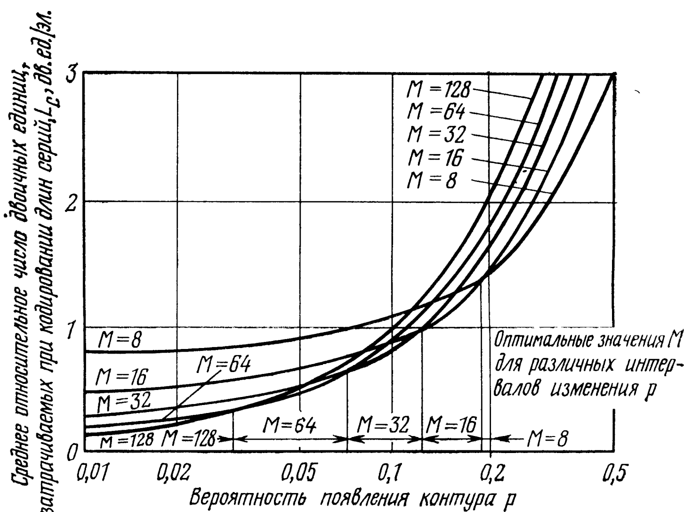


Рис. 22.3.3. Эффективность кодирования длин серий при передаче одноцветных изображений.

M — предел, установленный для длины серий; *A* — количество двоичных разрядов, отводимых на кодирование уровня серии.

0,1, то на передачу элемента изображения требуется в среднем менее 1 дв. ед., тогда как непосредственное применение ИКМ приводит в этом случае к затрате 3 дв. ед./эл.

Кодирование длин серий наиболее выгодно в случае изображений, допускающих грубую передачу полутонов. В случае факсимильного черно-белого изображения достаточно, указав предварительно уровень начального элемента строки, далее кодировать только расположение перепадов уровня. Относительная величина энтропии совпадает здесь с энтропией распределения вероятностей длин серий в соответствии с рис. 22.3.2. Если «черные» и «белые» серии при передаче черно-белых изображений существенно различаются по средней длине, то часто предпочтительно применить к ним различные коды. Коэффициент сокращения объема данных в факсимильных системах передачи черно-белых изображений с одномерным кодированием достигает по имеющимся сообщениям

величины 5 : 1. Обобщения метода кодирования длин серий, связанные с переходом к двум пространственным переменным [37—39], доводят это отношение до 10 : 1.

22.3.2. КОДИРОВАНИЕ РАЗРЯДНЫХ ПЛОСКОСТЕЙ

Изображение, закодированное b-разрядным равномерным кодом, можно представить в виде так называемых разрядных плоскостей, т. е. двоичных изображений, образованных одинаковыми разрядами кодовых слов. Оказалось, что в плоскостях старших разрядов значения элементов меняются редко, тогда как значения элементов «младших» плоскостей флуктуируют почти беспорядочно. Если применительно к этим изменениям воспользоваться методом кодирования длин серий отдельно для каждой плоскости, то можно, вообще говоря, ожидать сокращения полосы частот передаваемого сигнала. Однако эксперименты по моделированию подобных систем показали, что, как правило, не достигается даже двукратное сокращение полосы частот [40, 41].

22.4. КОДИРОВАНИЕ С ПОПОЛНЕНИЕМ КАДРОВ

Весьма значительный резерв сокращения полосы частот телевизионного сигнала заключен в межкадровой избыточности телевизионного сообщения. Передача большинства сцен сопровождается относительно небольшими изменениями в деталях изображения от одного кадра к другому. Таким образом, если передавать одни лишь изменения в изображении по отношению к некоторому первоначально выбранному кадру, то можно получить существенный выигрыш в полосе частот. Основные трудности здесь связаны с тем, что метод получения межкадрового разностного сигнала должен быть хорошо согласован с процессами телевизионной развертки и передачи, а оборудование для запоминания и обработки сигналов должно быть не слишком громоздким.

Если взять разности уровней соответственных элементов данного и соседнего кадров, то окажется, что уровень большинства элементов сохраняется неизменным. Это обстоятельство привело к разработке Маунтсом метода межкадрового кодирования, названного кодированием с пополнением кадров [42—45]. Блок-схема системы кодирования с пополнением кадров приведена на рис. 22.4.1. Каждый кадр, поступающий в процессе передачи от телевизионной камеры, преобразуется в массив восьмиразрядных двоичных чисел, причем первый массив некоторой последовательности кадров запоминается. Для каждого последующего кадра яркость элементов сравнивается с яркостями соответственных элементов зарегистрированного образца. Если обнаруживается значительное расхождение, то на месте соответствующего

элемента образца прежнее значение яркости заменяется новым, поступающим одновременно в буферный накопитель для дальнейшей передачи в канал связи. Необходимо, кроме того, определить, где именно в пределах строки произошло изменение яркости, и передать код соответствующей горизонтальной координаты. Первому элементу каждой строки также присваивается

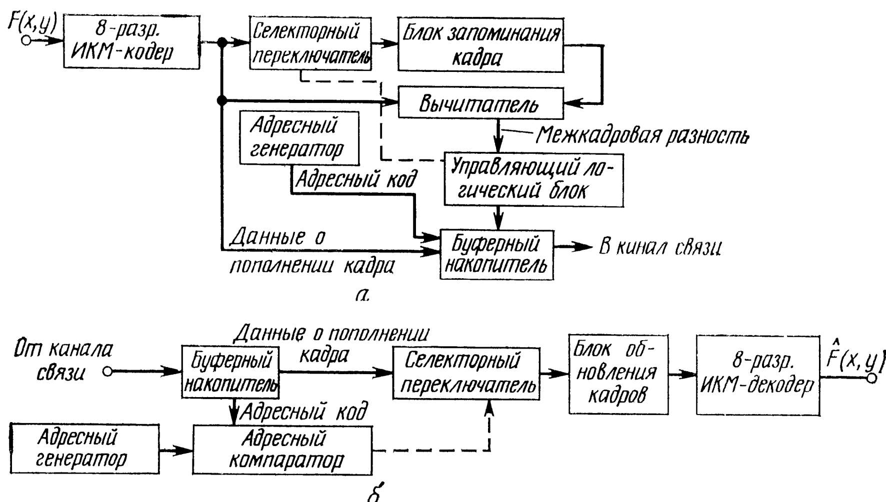


Рис. 22.4.1. Система кодирования изображений с пополнением кадров.
а — передатчик; *б* — приемник.

свой код, обеспечивающий счет числа строк. Дальнейшее увеличение эффективности системы может быть достигнуто при переходе от поэлементного кодирования межкадровых разностей к кластерному кодированию, учитывающему их тенденцию к пространственной группировке.

Одно из главных затруднений при построении систем с пополнением кадров связано с необходимостью буферного накопителя, емкость которого должна быть рассчитана на обработку длинных серий сигналов межкадровой разности, связанных с достаточно продолжительными движениями передаваемых объектов. Это требование может быть в некоторой степени смягчено путем регулирования порога чувствительности к межкадровым разностям в зависимости от степени заполнения буферного накопителя. Если накопитель почти пуст, то чувствительность может быть повышена с тем, чтобы пополнение кадров происходило более интенсивно, и наоборот, если накопитель близок к насыщению, то чувствительность может быть снижена. Объем накопителя можно также уменьшить, пространственно объединяя несколько элементов

в один отсчет на подвижных участках изображения. Эксперименты по моделированию системы с пополнением кадров [45] указывают на возможность снижения удельной затраты двоичных цифр до 1 дв. ед./эл.; качество телевизионной передачи при этом оценивается как отличное, за исключением периодов особо большой подвижности изображаемых объектов.

22.5. КОДИРОВАНИЕ С ПРЕДСКАЗАНИЕМ

Общая схема системы кодирования с предсказанием показана на рис. 22.5.1. В такой системе уровень каждого элемента изображения предсказывается на основании сведений о предшествующих по ходу развертки элементах. Оценка, полученная в результате предсказания, вычитается из фактической величины уровня, и разностный сигнал подвергается затем квантованию, кодированию и передаче по каналу связи. На приемной стороне по квантованному разностному сигналу восстанавливается видеосигнал. Сокращение полосы частот достигается за счет грубого квантования передаваемого разностного сигнала. Как показывает библиографический список, помещенный в конце этой главы, на разработку систем кодирования изображений с предсказанием были затрачены значительные усилия. В последующих разделах прослеживается весь путь, пройденный в этом направлении, с того времени, когда был выдвинут основной принцип, до современных исследований в области адаптивных систем кодирования с предсказанием.

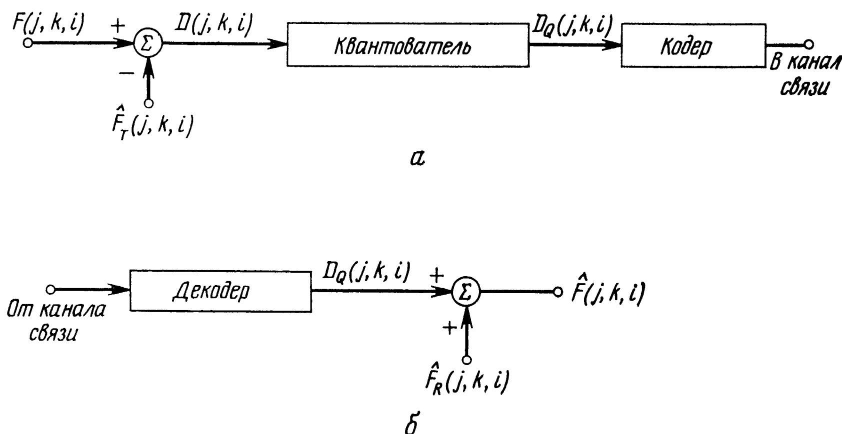


Рис. 22.5.1. Система кодирования с предсказанием.
а — схема кодирования; *б* — схема декодирования.

22.5.1. ДЕЛЬТА-МОДУЛЯЦИЯ

Система с дельта-модуляцией, показанная на рис. 22.5.2 [46], осуществляет кодирование с предсказанием в его простейшей форме. Аналоговый видеосигнал с ограниченной полосой частот подается на вход дифференцирующего (вычитающего) устройства. Если разностный сигнал на выходе устройства положителен, то в момент взятия отсчета от тактового генератора поступает положительный импульс стандартной амплитуды; в противном случае этот импульс будет отрицательным. Положительные и отрицательные импульсы передаются по каналу связи в виде двоичных цифр. Одновременно на передающей стороне последовательность импульсов подается в цепь обратной связи на интегратор, воспроизводящий входной сигнал в ступенчатой форме. На приемной стороне последовательность двоичных цифр преобразуется опять в положительные и отрицательные импульсы, интегрирование которых дает восстановленный видеосигнал. Фильтр нижних частот подавляет высокочастотные компоненты, внесенные в видеосигнал процессом кодирования. В этом варианте системы с дельта-модуляцией предсказание основывается на учете только одного предшествующего элемента, а разностный сигнал для каждого элемента квантуется всего лишь на два уровня.

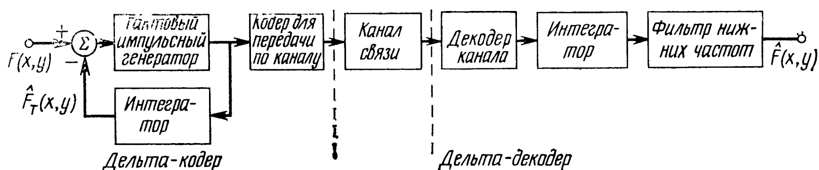


Рис. 22.5.2. Система кодирования с дельта-модуляцией.

Система дельта-модуляции с двумя уровнями квантования для кодирования сигналов разработана исследовательскими лабораториями фирмы «Филипс» (Голландия, 1952 г.) [47]. В 1961 г. Саламан [48] описал систему дельта-модуляции, предназначенную для передачи телевизионных сигналов. Одна из первых экспериментальных систем телевизионной передачи с дельта-модуляцией была разработана Бальдером и Крамером [49]. В этой системе прошли испытания интегрирующие цепи двух типов. Простейшая из них представляла собой одиночную RC-цепь, другая состояла из нескольких последовательно соединенных RC-цепей с различными постоянными времени. Полученные при этом результаты достаточно показательны с точки зрения оценки возможностей простейших систем дельта-модуляции с двумя уровнями квантования. В варианте с одиночной интегрирующей RC-цепью частота

отсчетов была установлена с 10-кратным превышением частоты Найквиста. Поскольку дельта-модуляция предполагает затрату 1 дв. ед. на отсчет, для передачи требовалось 10 дв. ед. в пересчете на каждый из элементов изображения, взятых с интервалом Найквиста. Это больше, чем 6—8 дв. ед., которые затрачивает на передачу элемента обычная система ИКМ, так что одиночная интегрирующая цепь приводила к расширению полосы частот. В системе с комбинированной интегрирующей цепью для сохра-

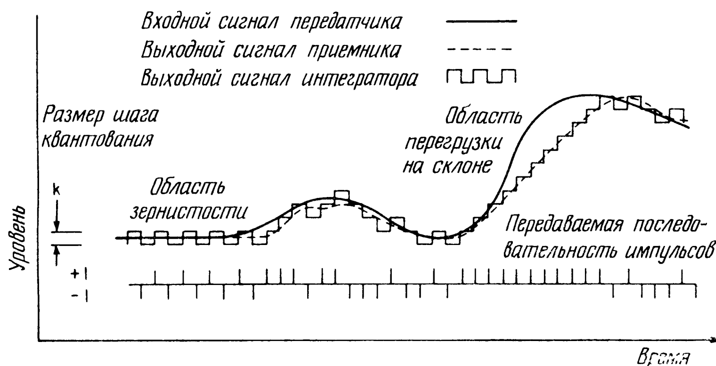


Рис. 22.5.3. Структура сигналов в системе кодирования с дельта-модуляцией.

нения того же качества передачи требовалось 4-кратное превышение частоты Найквиста, что сводило расход двоичных цифр к 4 дв. ед./эл.

Рис. 22.5.3 иллюстрирует характерные особенности работы системы с дельта-модуляцией. Чертеж поясняет основную трудность в построении таких систем: если шаг квантования выбирается достаточно малым, чтобы ограничить шум квантования, то крутые склоны видеосигнала удается воспроизвести достаточно точно не иначе, как ценой резкого повышения частоты отсчетов. В случае фиксированной частоты отсчетов остается, таким образом, лишь возможность обменивать уровень шума квантования на степень «перегрузки» системы на склонах.

На рис. 22.5.4 приведены фотоснимки, полученные путем цифрового моделирования процесса дельта-модуляции. В случае когда шаг ступени при воспроизведении яркости составляет 2,5 % от пикового уровня яркости изображения, кодирование сопровождается заметным эффектом перегрузки на склонах. Если же взять другой крайний случай и установить шаг ступени равным 10 % от пикового значения яркости, то на первый план выступает шум, проявляющийся как зернистость изображения. Шаг с относительной величиной 5 % дает приемлемый компромисс между перегрузкой и зернистостью.

Действие кодера в системе дельта-модуляции сводится по существу к дифференцированию видеосигнала, вследствие чего при декодировании возникает необходимость в интегрирующем звене

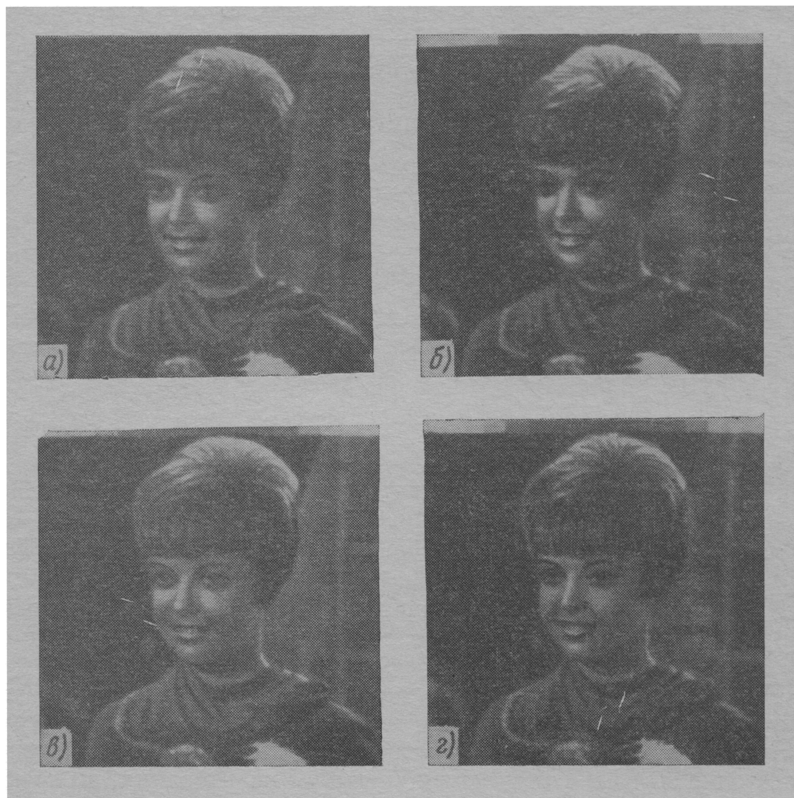


Рис. 22.5.4. Примеры кодирования изображений посредством дельта-модуляции с затратой 1 дв. ед./эл.

q — шаг ступенчатого воспроизведения яркости; СКО — среднеквадратическая ошибка воспроизведения яркости

а) оригинал; *б*) $q = 2{,}5$ %, СКО = 5,71 %; *в*) $q = 5$ %, СКО = 2,58 %; *г*) $q = 10$ %, СКО = 3,35 %.

для восстановления среднего уровня яркости исходного изображения. Поскольку интегрирование дельта-импульсов производится после того, как они переданы по каналу связи, возникающие при передаче ошибки ведут к накоплению погрешности воспроизведения яркости в пределах строки. В качестве средства, позволяющего избежать интегрирования на приемной стороне, Инозе и Ясуда [50] предложили систему кодирования с дельта-сигма-модуляцией, показанную на рис. 22.5.5. Как показано на

рис. 22.5.5, *а*, интегратор помещен перед дельта-кодером, а дифференцирующее звено следует за ним. Эта схема, иллюстрирующая общий принцип дельта-сигма-модуляции, может быть сведена к более простой схеме рис. 22.5.5, *б* путем некоторой перестановки линейных звеньев системы. Показано, что дельта-сигма-модулятор представляет собой особую форму одноразрядного кодера с непосредственной обратной связью [51].

Чтобы подавить эффект перегрузки на склонах, избежав при этом увеличения шума квантования, можно, в частности, перейти

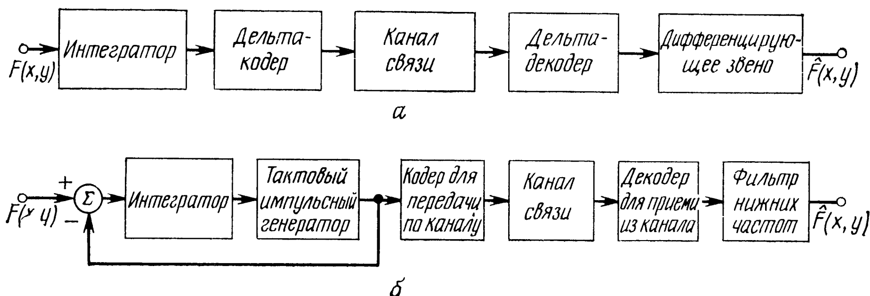


Рис. 22.5.5. Система кодирования с дельта-сигма-модуляцией для передачи изображений.
а — общий принцип построения системы; *б* — упрощенная система.

к системе дельта-модуляции, где допустимы уже не два, а большее число уровней квантования. Действительно, качество работы многоуровневых дельта-модуляторов повышается, однако достигается это ценой дополнительных усложнений. Следует при этом отметить, что многоуровневая дельта-модуляция представляет собой в действительности частный случай дифференциальной импульсно-кодовой модуляции, рассматриваемой ниже. Другой путь повышения качества работы дельта-модулятора состоит в адаптации дискретной шкалы яркостей по отношению к параметрам видеосигнала в процессе воспроизведения изображения. Адаптивные линейные системы с предсказанием будут рассмотрены в следующем разделе.

22.5.2. ДИФФЕРЕНЦИАЛЬНАЯ ИМПУЛЬСНО-КОДОВАЯ МОДУЛЯЦИЯ

Исходным пунктом для развития общей концепции линейного адаптивного кодирования послужило изобретение Катлером [52] системы дифференциальной импульсно-кодовой модуляции (ДИКМ). Катлер предложил применить интеграторы для предсказания уровня данного элемента на основании уровня предшествующего элемента строки, а разность между оценкой, полу-

ченной в результате предсказания, и фактической величиной уровня подвергать квантованию и кодированию для передачи по каналу связи.

Блок-схема системы кодирования изображений с ДИКМ приведена на рис. 22.5.6. В системе такого типа непрерывный сигнал, соответствующий изображению, дискретизируется по пространственным переменным, и разность между фактическим уровнем

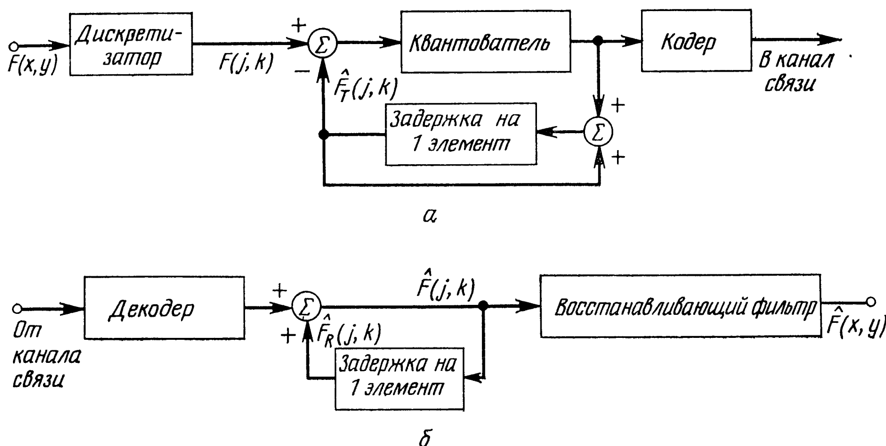


Рис. 22.5.6. Система ДИКМ для передачи изображений.
а — передатчик; *б* — приемник.

элемента и его оценкой квантуется и кодируется для передачи. Для квантования разности используют обычно 8 уровней с дальнейшим преобразованием в трех- или четырехразрядные двоичные комбинации [53, 54]. Таким образом, сокращение полосы частот определяется разницей между 6—8 дв. ед. и 3 дв. ед., затрачиваемыми на передачу элемента изображения соответственно с помощью ИКМ и ДИКМ. В типичном кодере системы ДИКМ предсказание для данного элемента основано на квантованном разностном сигнале, полученном для предшествующего элемента строки. На приемной стороне декодированный и восстановленный разностный сигнал комбинируется со своим приближением, полученным в результате предсказания точно тем же способом, что и на передающей стороне, и таким образом восстанавливается исходное изображение.

В 1958 г. Грэхэм [55] предложил шкалу квантования с неравномерным расположением уровней. Представление о типичных шкалах такого рода, предназначенных для систем ДИКМ и дельта-модуляции, дает рис. 22.5.7. Применение неравномерной шкалы существенно повышает субъективную оценку качества восстанав-

ливаемого изображения. Однако, как показали эксперименты, для большинства приложений при этом приходится сохранять по меньшей мере 8 уровней квантования.

Входной сигнал изменяется в определенном диапазоне уровней яркости. Следовательно, предсказанный сигнал также не должен выходить за пределы этого диапазона. Если уровень входного отсчета достигает верхней границы диапазона, то аналоговый разностный сигнал для данного элемента может принимать либо нулевое, либо отрицательное значение. В другом крайнем случае, когда входной отсчет минимален, разностный сигнал может иметь либо нулевое, либо положительное значение. В любой из этих ситуаций некоторая часть всей биполярной шкалы квантования заведомо не понадобится. Мусман [56] в связи с этим предложил «перекидную» шкалу квантования, идею которой поясняет рис. 22.5.8. Такая шкала позволяет использовать положительные (отрицательные) уровни квантования, когда они заведомо не нужны в ее нижней (верхней) части. Хорошо спроектированная восьмиуровневая (трехразрядная) система ДИКМ по субъективной оценке качества получаемых изображений в общем приближается к пяти-шестиразрядной системе ИКМ; отмечаются лишь небольшие искажения вблизи резких контуров [57]. При демонстрации изображений в виде последовательности кадров в реальном масштабе времени телевизионной передачи определенная степень корреляции таких искажений на соседних кадрах вызывает специфический эффект, называемый «контурной активностью» (edge busyness). Эта активность проявляется в виде искрящегося обрамления контурных участков изображения.

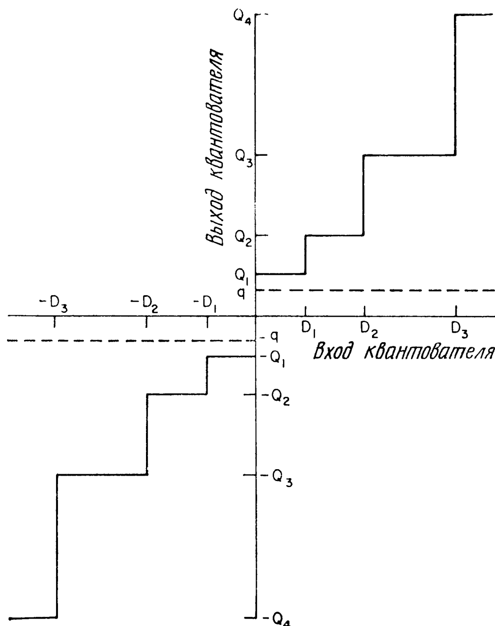


Рис. 22.5.7. Шкала квантования для систем дельта-модуляции и ДИКМ.

Квантованные разностные сигналы, формируемые кодером в восьмиуровневой системе ДИКМ, различаются вероятностью

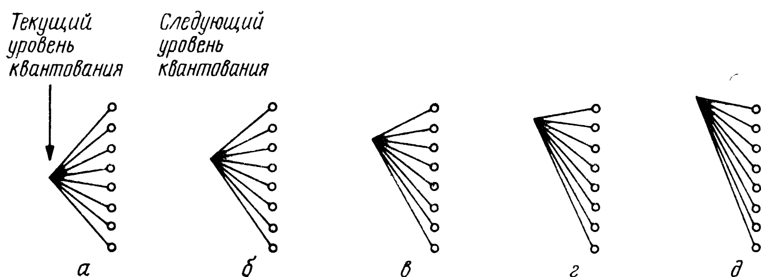


Рис. 22.5.8. «Перекидная» шкала квантования для системы ДИКМ.
а — вдали от верхней границы диапазона яркости; *б* — на расстоянии трех элементов от границы; *в* — на расстоянии двух элементов от границы; *г* — на расстоянии одного элемента от границы; *д* — на границе.

Рис. 22.5.9. Примеры кодирования изображений посредством ДИКМ с затратой 3 дв. ед./эл.
Q_1 — наименьший уровень квантования; СКО — среднеквадратическая ошибка воспроизведения яркости.
а) оригинал; *б*) $Q_1 = 0{,}5$ %, СКО = 1,31 %; *в*) $Q_1 = 1{,}0$ %, СКО = 1,33 %; *г*) $Q_1 = 5{,}0$ % СКО = 1,66 %.

своего появления — более многочисленны малые разности. Можно, следовательно, вместо трехразрядного равномерного кода применить статистический код с переменной длиной слова, например код Хаффмэна, что приводит к дальнейшему сокращению объема передаваемых данных. Как показывает моделирование, код Хаффмэна снижает удельный расход двоичных цифр до 2,5 дв. ед./эл. [58, 59]. Однако применение такого кода осложняется необходимостью использования буферного накопителя данных.

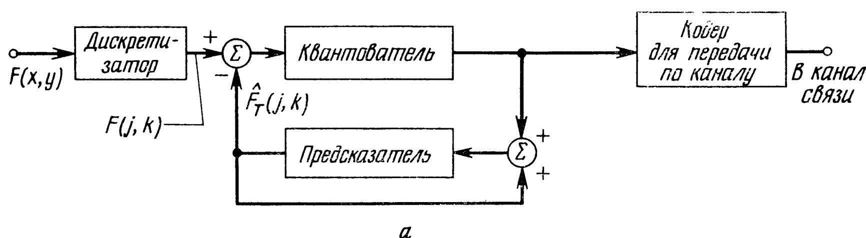
На рис. 22.5.9 приведены фотоснимки, полученные в результате цифрового моделирования системы ДИКМ. В экспериментах использовалась неравномерная шкала квантования типа показанной на фиг. 22.5.7 в трех вариантах, различавшихся величиной наименьшего уровня квантования Q_1. По субъективным оценкам, при $Q_1 = 2{,}5$ % достигается наилучший компромисс между искажениями двух видов — перегрузкой и зернистостью.

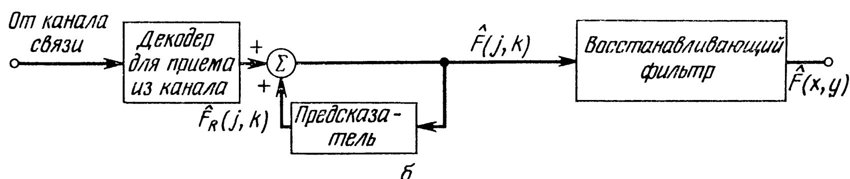
Брэйнард и Кэнди [51, 60] провели исследования по применению кодера с непосредственной обратной связью для передачи изображений. Такой кодер, будучи упрощенным вариантом предложенного Киммом и Кью [61] кодера с шумовой обратной связью, содержит дифференцирующий фильтр для предварительного обострения перепадов сигнала, интегрирующий усилитель и сглаживающий (путем интегрирования с утечкой) оконечный фильтр. Предполагалось, что по сравнению с кодером системы ДИКМ преимущество кодера с непосредственной обратной связью будет определяться большей свободой его конструирования в связи с возможностью подбора частотных характеристик трех составляющих его фильтров таким образом, чтобы свести к минимуму шум квантования. Однако при моделировании такой системы [51] не обнаружилось сколько-нибудь значительного преимущества по сравнению с хорошо спроектированными кодерами, использующими неравномерную шкалу квантования.

22.5.3. КОДЕРЫ С ПРОСТРАНСТВЕННЫМ ПРЕДСКАЗАНИЕМ

В 1952 г. Гаррисон [62] расширил концепцию дифференциальной импульсно-кодовой модуляции, предложив основывать предсказание на линейной комбинации уровней нескольких элементов, принадлежащих как пройденному участку строки, так и предшествующим строкам. Теоретические и экспериментальные исследования показали, что предсказание на основе более полных сведений дает возможность улучшить качество изображений как при объективной оценке среднеквадратической ошибкой воспроизведения, так и при визуальной оценке [63, 64]. Обобщенная блок-схема системы кодирования с пространственным предсказанием представлена на рис. 22.5.10; там же для удобства дальнейшего рассмотрения дана нумерация элементов, используемых

в процессе предсказания. Схема предсказателя будет более подробно описана в разд. 22.6. Стандартный кодер системы ДИКМ, предсказание которого для элемента S_0 основано на учете одного предшествующего элемента S_1 той же строки, часто характери-





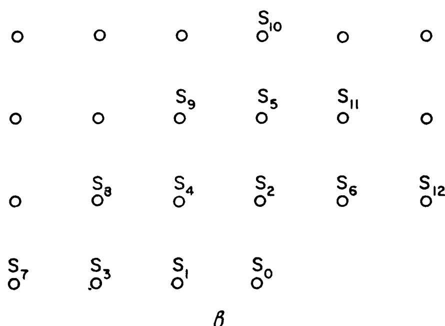


Рис. 22.5.10. Система кодирования с пространственным предсказанием для передачи изображений.

а — передатчик; *б* — приемник; *в* — нумерация элементов изображения.

зуют как предсказатель первого порядка. В соответствии с такой классификацией предсказатель второго порядка будет учитывать два предыдущих элемента строки (S_1 и S_3) или, возможно, предшествующий элемент строки (S_1) и ближайший элемент предыдущей строки (S_2). Предсказатель третьего порядка может использовать для предсказания тройку элементов (S_1, S_2, S_4). Для предсказания часто используется также элемент S_6, поскольку он приобретает важную роль для передачи контуров вертикальной ориентации. Как показали исследования, более подробно описы-

ваемые в разд. 22.6, для уменьшения искажений при кодировании наиболее существенны данные о тех близлежащих элементах, по отношению к которым статистические измерения указывают на максимальную степень корреляции с оцениваемым элементом. Установлено также, что использование порядка предсказания выше третьего уже не дает ощутимого дальнейшего снижения погрешностей кодирования.

Основная трудность, связанная с повышением порядка линейного предсказания при кодировании изображений, состоит в усложнении кодирующей системы. Например, для системы с предсказанием третьего порядка, обращающейся при оценке элемента S_0 к элементам на позициях S_1, S_2, S_4, требуется запоминание трех квантованных разностей, а именно $S_1 - \hat{S}_1$, $S_2 - \hat{S}_2$, $S_4 - \hat{S}_4$. Кроме того, сам предсказатель получается в этом случае несколько более сложным по сравнению с кодером системы ДИКМ.

22.5.4. АДАПТИВНЫЕ КОДЕРЫ С ПРЕДСКАЗАНИЕМ

Как было отмечено выше, применение дельта-модуляции для кодирования изображений связано с поиском компромисса между искажениями двух типов: зернистостью и перегрузкой на склонах видеосигнала. Оба этих эффекта могут быть в определенной степени ослаблены путем адаптации шкалы квантования к характеристикам изображения в процессе кодирования. В случае быстрых изменений яркости вдоль строки развертки выгодно увеличить шаг квантования — это ослабит перегрузку на склонах. Если же яркость изменяется очень медленно, шаг квантования выгодно уменьшить для устранения зернистости. Такой подход лежит в основе ряда адаптивных систем кодирования с дельта-модуляцией и с ДИКМ [65—74]. Алгоритм квантования для кодирования с дельта-модуляцией, предложенный Винклером [65, 66], предусматривает удвоение шага квантования в том случае, когда зарегистрированы три дельта-сигнала одной полярности, следующие подряд, и снижение шага вдвое всякий раз, когда дельта-сигналы меняют знак. Теоретическое исследование этой процедуры привело Джайэнта [75] к заключению, что увеличение шага квантования следует производить в отношении $^3/_2 : 1$, а уменьшение — в отношении $^2/_3 : 1$. В принципе можно, конечно, распространить этот подход на случай ДИКМ, рассматривая возможность адаптивной расстановки всех восьми уровней квантования. Однако построение такого кодера затрудняется неизбежным дополнительным его усложнением, а также опасностью потери устойчивости системой в целом.

Другой подход к адаптивному кодированию с предсказанием заключается в двухрежимном кодировании. Кодер такой системы работает либо в режиме ДИКМ, либо в режиме дельта-модуляции.

Если на протяжении нескольких отсчетов яркость изображения меняется достаточно плавно, то кодирование ведется в режиме дельта-модуляции; резкие изменения яркости вызывают переключение кодера в режим трехразрядной ДИКМ до тех пор, пока видеосигнал опять не примет более спокойный характер. Достаточно простой алгоритм переключения режимов состоит в том,

Таблица 22.5.1

Логическая схема образования кодовых комбинаций в двухрежимной системе с избыточной дискретизацией

Режим дельта-модуляции			Передаваемый код					
+q	+q	+q	1	1	1	(1	0)	← Маркерные разряды
+q	+q	−q	1					
+q	−q	−q	1					
−q	+q	+q	1					
−q	−q	+q	0					
−q	+q	−q	0					
+q	−q	−q	0					
−q	−q	−q	0	0	0	(0	1)	← Маркерные разряды

После трех „единиц" подряд, обнаруженных в последовательности дельта-сигналов, передается одноразрядный маркерный знак „нуль"
···1 1 1 0···

После трех „нулей" подряд, обнаруженных в последовательности дельта-сигналов, передается одноразрядный маркерный знак „единица"
···0 0 0 1···

чтобы переходить от дельта-модуляции к ДИКМ после трех идентичных дельта-сигналов, следующих подряд. Сигнал для обратного переключения от ДИКМ к дельта-модуляции возникает в том случае, когда кодер работает в сущности вхолостую, используя свою 8-уровневую шкалу квантования лишь для переходов между уровнями $-Q_1$ и $+Q_1$. Основным преимуществом такого алгоритма является его простота: он не требует передачи специальных дополнительных сигналов для синхронизации переключений на приемной стороне. К сожалению, платой за это преимущество является запаздывание с переключением от дельта-модуляции к ДИКМ на три элемента, что может приводить к ощутимому эффекту перегрузки и соответствующим искажениям в передаче контуров. Более тонкая процедура двухрежимного кодирования была предложена Фраем, Шиндлером и Веттигером [74]. Чтобы повысить точность в определении момента перехода от дельта-модуляции к ДИКМ, здесь предусматривается избыточная дискретизация изображения с увеличением в 3 раза нормаль-

ной частоты отсчетов. Процесс переключения системы из режима дельта-модуляции в режим ДИКМ иллюстрирует табл. 22.5.1. Если в режиме дельта-модуляции, когда дискретизация избыточна, зарегистрированы три одинаковых дельта-сигнала подряд, передается пятиразрядная двоичная кодовая комбинация в соответствии с табл. 22.5.1. В остальных случаях для каждой тройки дельта-сигналов решение принимается «большинством голосов»:



Рис. 22.5.11. Примеры адаптивного кодирования изображений посредством линейного предсказания.

а — двухрежимное кодирование с затратой 2 дв ед /эл ; *б* — двухрежимное кодирование с избыточной дискретизацией с затратой 2 дв ед /эл.

передается «единица», если два сигнала в тройке положительны, в противном случае передается «нуль». Возвращение от ДИКМ к дельта-модуляции производится при возникновении холостого хода, т. е. при прямом или обратном переходе от уровня $-Q_1$ к уровню $+Q_1$ в шкале квантования.

Как показали эксперименты по изучению работы двухрежимного кодера в реальной телевизионной системе, такое же субъективное качество передачи, что и в трехразрядной системе ДИКМ, удается обеспечить при затрате в среднем около 2,0 дв. ед./эл. Метод имеет и недостатки, связанные с усложнением процессов кодирования и декодирования, а также с повышением по сравнению с системой ИКМ чувствительности к шумам канала связи. Кроме того, переменная длина кодовых слов ведет к необходимости буферного накопления данных.

На рис. 22.5.11 приведены фотоснимки, полученные путем цифрового моделирования системы двухрежимного кодирования в простом ее выполнении и с избыточной дискретизацией. По субъективным оценкам, двухрежимное кодирование дает значительно лучший результат, чем адаптивное кодирование с дельта-

модуляцией, однако удваивает расход двоичных цифр. Простые системы двухрежимного кодирования не освобождают в полной мере от искажений типа перегрузки вблизи контуров изображения. Заметность таких искажений в двухрежимных системах с избыточной дискретизацией оказывается значительно меньшей.

22.5.5. КОДИРОВАНИЕ С ПРЕДСКАЗАНИЕМ ПРИМЕНИТЕЛЬНО К ЦВЕТНЫМ ИЗОБРАЖЕНИЯМ

Для применения кодирования с предсказанием в системах цветного телевидения имеются в принципе два пути: непосредственное кодирование, объектом которого служит полный цветной телевизионный сигнал, и кодирование координат цвета. Возможность кодирования с дельта-модуляцией применительно к полному сигналу системы NTSC была изучена Тичером и Ютцем [76], которые установили, что недостаточная чувствительность системы с дельта-модуляцией к высокочастотным составляющим сигнала приводит к неприемлемым амплитудным и фазовым искажениям поднесущей сигнала цветности, если не увеличивать чрезмерно частоту дискретизации. Возможность использования дельта-модуляции для кодирования трех координат цвета R, G, B при телевизионной передаче изучалась Бишаном [77]. Лимб, Рубинштейн и Уолш [78] исследовали эффективность применения ДИКМ для кодирования сигналов яркости и цветности

$$Y = 0{,}299R_N + 0{,}587G_N + 0{,}114B_N, \qquad (22.5.1а)$$

$$C_1 = (R_N - Y)\frac{\cos\theta}{1{,}14} - (B_N - Y)\frac{\sin\theta}{2{,}03}, \qquad (22.5.1б)$$

$$C_2 = (R_N - Y)\frac{\sin\theta}{1{,}14} + (B_N - Y)\frac{\cos\theta}{2{,}03}, \qquad (22.5.1в)$$

определяемых координатами цвета R_N, G_N, B_N в системе координат телевизионного приемника NTSC, а также параметром θ преобразования координат цвета. Для достижения приемлемого качества видеотелефонного изображения потребовалось положить $\theta = 22°$, а для квантования разностных сигналов ДИКМ в каналах Y, C_1 и C_2 установить соответственно 12, 6 и 4 уровня. Кроме того, сигналы цветности C_1 и C_2 кодировались с пропуском каждой второй строки развертки; утерянные отрезки сигналов цветности восстанавливались при декодировании путем усреднения строк. Подходя к проблеме несколько иначе, Хабиби [79] предложил схему кодирования, согласно которой сигнал яркости изображения кодируется с применением ДИКМ по стандартной методике, а сигналы цветности I и Q системы NTSC кодируются каждый в отдельности с применением адаптивной дельта-модуляции.

22.5.6. ДИФФЕРЕНЦИАЛЬНОЕ МЕЖКАДРОВОЕ КОДИРОВАНИЕ

Кодирование с предсказанием можно распространить на последовательность кадров, чтобы использовать наличие межкадровой избыточности [45]. Как правило, межкадровая избыточность достигает большой величины, так что кодирование с межкадровым предсказанием могло бы оказаться эффективным. Тем не менее исследования и разработки в этой области сравнительно малочисленны, что можно объяснить практическими трудностями запоминания последовательностей кадров. Маунтс [80] осуществил межкадровую ДИКМ в виде системы кодирования с предсказанием, при котором учитывается лишь один предшествующий кадр. Затрачивая на передачу элемента 3 дв. ед., такая система по качеству работы примерно соответствует ДИКМ с предсказанием по предшествующему элементу строки. Кодер в системе межкадровой ДИКМ может испытывать «временну́ю перегрузку», аналогичную перегрузке на склонах в обычной системе ДИКМ. Искажениям за счет временно́й перегрузки особенно подвержены крупные, высококонтрастные подвижные объекты; при их воспроизведении контуры теряют резкость и тем заметнее, чем меньше максимальный уровень квантования. Испытывались также возможности совместного применения межкадровой ДИКМ и метода пополнения кадров, описанного в разд. 22.4 [80].

Работа систем ДИКМ с пространственно-временны́м предсказанием изучалась путем цифрового моделирования [81]. Представляется возможным обеспечить хорошее качество воспроизведения при передаче с затратой в среднем около 2 дв. ед. на элемент при наличии, разумеется, средств для запоминания кадров.

22.6. ПРОЕКТИРОВАНИЕ СИСТЕМ ЛИНЕЙНОГО ПРЕДСКАЗАНИЯ ДЛЯ КОДИРОВАНИЯ ИЗОБРАЖЕНИЙ

В процессе проектирования системы линейного предсказания необходимо определить следующие параметры кодера: количество и расположение пройденных в ходе развертки элементов, уровни которых будут использоваться для предсказания, весовой коэффициент для каждого из этих элементов, а также количество и расположение уровней квантования. Нелинейный характер операции квантования создает значительные аналитические трудности для оптимизации конструкции кодера с целью обеспечения минимальной ошибки воспроизведения изображения.

Большинство методов расчета предусматривает оптимизацию предсказателя без учета ошибок квантования, а затем оптимизацию, снижающую до минимума ошибку квантования [82—86]. Ниже мы будем следовать именно такому подходу.

При кодировании изображений в системе с предсказанием (рис. 22.6.1) предсказатель вырабатывает оценку $\hat{F}_T(j, k, i)$ фактического уровня $F(j, k, i)$ данного элемента на основе сведений об уровнях предшествующих по ходу развертки элементов, принадлежащих тому же кадру или предшествующим кадрам.

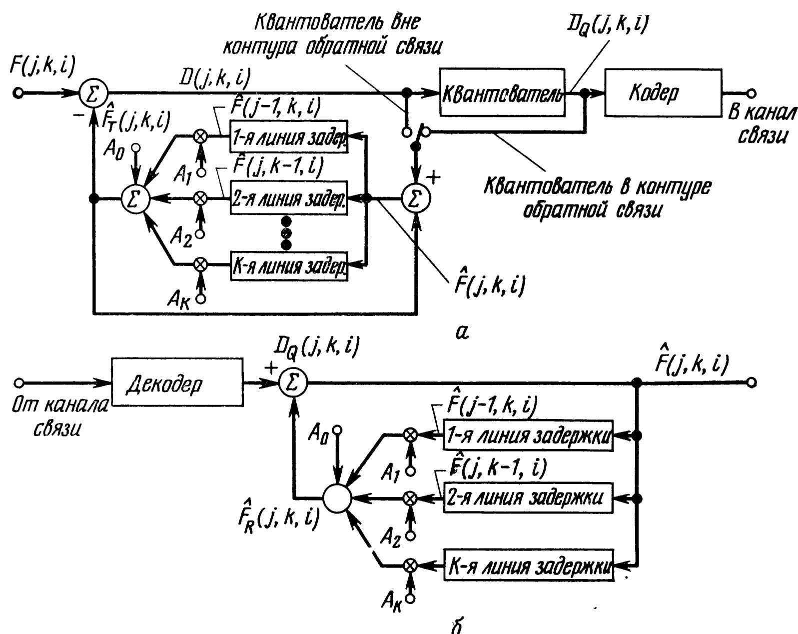


Рис. 22.6.1. Система кодирования с линейным предсказанием по замкнутой схеме для передачи изображений.
а — кодирующий блок; *б* — декодирующий блок.

Ошибка предсказания $D(j, k, i)$ — разность оценки уровня и его фактической величины для данного элемента

$$D(j, k, i) = F(j, k, i) - \hat{F}_T(j, k, i) \tag{22.6.1}$$

— квантуется, кодируется и передается по каналу связи. На приемной стороне декодированная квантованная ошибка предсказания комбинируется с оценкой уровня элемента для восстановления значения уровня:

$$\hat{F}(j, k, i) = D_Q(j, k, i) + \hat{F}_R(j, k, i). \tag{22.6.2}$$

Если удалить квантователь из схемы рис. 22.6.1, то восстановленный уровень $\hat{F}(j, k, i)$ элемента совпадет с его фактической

величиной $F(j, k, i)$ [без учета погрешностей оценки $\hat{F}_R(j, k, i)$]. Требования к пропускной способности канала связи могут быть снижены лишь путем квантования и кодирования ошибки предсказания. Обычно для квантования этой величины используют от 2 до 16 уровней, а для кодирования — от 1 до 4 двоичных разрядов.

Для оценки $\hat{F}_T(j, k, i)$ уровня данного элемента в передатчике теоретически имеются уровни всех пройденных в процессе развертки элементов данного и предшествующего кадров, а также квантованные ошибки предсказания $D_Q(j, k, i)$ для каждого из этих элементов. Приемник системы располагает восстановленными значениями уровней всех предшествующих элементов и квантованными ошибками предсказания для каждого из этих элементов.

Возможны два принципиально различных метода предсказания, которые следует рассмотреть с точки зрения их применимости в кодирующих системах. Один из этих методов — прямое предсказание, по разомкнутой схеме (feedforward), а другой — предсказание с обратной связью, по замкнутой схеме (feedback). Система с предсказанием по разомкнутой схеме строит оценку в виде линейной комбинации уровней предшествующих элементов, обращаясь за этими данными к запоминающему устройству. Основная трудность при построении систем с прямым предсказанием состоит в том, что декодирующий блок системы в отличие от кодирующего блока в этом случае не имеет доступа к зарегистрированной в запоминающем блоке совокупности фактических уровней предыдущих элементов; декодер может оперировать только восстановленными уровнями, значения которых в результате квантования определены с погрешностью. Таким образом, оценки $\hat{F}_T$ и $\hat{F}_R$ уровня данного элемента, полученные в передатчике и приемнике, будут различаться между собой, что в конечном счете приведет систему к потере устойчивости. Эта трудность может быть устранена предсказанием с обратной связью.

В системе кодирования изображений с предсказанием, построенной по замкнутой схеме (рис. 22.6.1), в передатчике получается оценка в виде линейной комбинации прошлых ошибок предсказания. Выходной сигнал первой линии задержки соответствует ошибке предсказания для соседнего элемента слева, а выходной сигнал второй линии задержки — для соседнего элемента сверху. Весовые коэффициенты A_k обычно подбираются таким образом, чтобы минимизировать дисперсию ошибки предсказания. Если квантователь не включен в контур обратной связи, то сложение оценки $\hat{F}_T$ для предшествующего элемента с разностным сигналом D приводит к точному воспроизведению уровня очередного

элемента с задержкой на один элемент. Затем производятся квантование и кодирование разностного сигнала для передачи по каналу связи. На приемной стороне квантованная ошибка D_Q прибавляется к оценке $\hat{F}_R$ приемника; полученная сумма есть восстановленное значение $\hat{F}$ уровня элемента. Если бы разностный сигнал не подвергался квантованию, то оценки передатчика и приемника были бы идентичны и восстановленное значение $\hat{F}$ представляло бы собой задержанный на межэлементный интервал уровень F очередного элемента. Квантование разчостного сигнала приводит к расхождению между $\hat{F}$ и F при восстановлении уровня. Это расхождение может накапливаться, поскольку оценки передатчика и приемника будут неодинаковы, если квантователь не включен в контур обратной связи. Накопление ошибок можно устранить, включив квантователь в состав контура обратной связи. В результате передатчик и приемник будут давать одинаковые оценки, поскольку в этом случае работу предсказателей как передатчика, так и приемника будет определять одинаковый квантованный разностный сигнал. Контур обратной связи, построенный по такой схеме, может быть с одинаковым успехом спроектирован в расчете на достижение минимума как ошибки квантования, так и дисперсии разностного сигнала.

Рассмотрим теперь вопрос об определении весовых коэффициентов предсказателя. Пусть в соответствии с рис. 22.5.10 S_1, S_2, ..., S_K обозначают уровни K предшествующих элементов, используемых для предсказания уровня элемента с фактическим уровнем S_0. Предшествующие элементы могут находиться на той же строке и на предыдущих строках. При расчете кодера уровни S_k этих элементов рассматриваются как отсчеты некоторой реализации случайного процесса с известными моментами первого и второго порядка.

Если пренебречь операцией квантования, то линейная оценка уровня элемента S_0 определяется как

$$\hat{S}_0 = A_0 + A_1 S_1 + A_2 S_2 + \cdots + A_K S_K, \tag{22.6.3}$$

где A_k — весовые коэффициенты предсказания. Ошибка предсказания

$$D = S_0 - \hat{S}_0. \tag{22.6.4}$$

В большинстве случаев весовые коэффициенты предсказания выбираются так, чтобы минимизировать среднеквадратическую ошибку предсказания

$$D_P \equiv E\{D^2\} = E\{[S_0 - \hat{S}_0]^2\}. \tag{22.6.5}$$

Преимущество такого критерия для оценки качества работы состоит в том, что он удобен с точки зрения анализа, достаточно хорошо согласуется с субъективной оценкой качества и, как подтвердится в дальнейшем, оказывается в прямой пропорциональной зависимости от ошибки квантования. Для минимизации среднеквадратической ошибки предсказания следует вычислить частные производные функции D_P по каждой из переменных A_k и приравнять каждую производную нулю. Итак, пусть

$$\partial D_P/\partial A_k = \partial E\,\{[S_0 - (A_0 + A_1S_1 + A_2S_2 + \cdots + A_KS_K)]^2\}/\partial A_k = 0. \tag{22.6.6}$$

Дифференцирование дает следующий результат:

$$E\,\{S_0\} = E\,\{\hat{S}_0\} = A_0 + A_1E\,\{S_1\} + A_2E\,\{S_2\} + \cdots + A_KE\,\{S_K\}, \tag{22.6.7а}$$

$$E\,\{S_0S_k\} = A_1E\,\{S_1S_k\} + A_2E\,\{S_2S_k\} + \cdots + A_KE\,\{S_KS_k\} \tag{22.6.7б}$$

при $k = 1, 2, \ldots, K$. Совместное решение $K + 1$ уравнений (22.6.7) определяет оптимальные весовые коэффициенты A_k через математические ожидания $E\,\{S_k\}$ всех элементов и смешанные моменты $E\{S_jS_k\}$ всех возможных пар элементов, используемых для предсказания. Если считать, что переменные A_k имеют оптимальные значения, то минимум среднеквадратической ошибки предсказания определяется выражением

$$(\sigma_P^2)_{\text{мин}} = E\,\{S_0^2\} - A_0E\,\{S_0\} - \sum_{k=1}^{K} A_k\,\{S_0S_k\}. \tag{22.6.8}$$

Известно, что в общем случае при неограниченном увеличении K достигается полная декорреляция совокупности ошибок предсказания [87]. Если же изображение представляет собой реализацию случайного процесса, характеризующегося авторегрессией N-го порядка, то для полной декорреляции требуется только N предшествующих элементов [86]. В этом случае дальнейшее увеличение числа элементов уже не повышает точности предсказания.

В табл. 22.6.1 приведен список весовых коэффициентов предсказателя для случая двумерного марковского процесса с автокорреляционной функцией

$$R\,(j_1,\ j_2;\ k_1,\ k_2) = \exp\left[-\sqrt{\alpha_C\,(j_1 - j_2)^2 + \alpha_R\,(k_1 - k_2)^2}\right], \tag{22.6.9}$$

где k_i — номера строк, j_i — номера столбцов, а α_R и α_C — масштабные коэффициенты в направлении строк и столбцов.

Таблица 22.6.1

Весовые коэффициенты и среднеквадратическая ошибка предсказания для различных вариантов кодирующих систем с линейным предсказанием. Указаны порядок предсказания и характеристика относительного расположения элементов, используемых для предсказания.
$\alpha_C = 16{,}7\cdot 10^{-4}$, $\alpha_R = 9{,}28\cdot 10^{-4}$, $\exp(-\alpha_C) = 0{,}960$, $\exp(-\alpha_R) = 0{,}970$

Весовые коэффициенты	1-го порядка, гор.	1-го порядка, верт.	2-го порядка, гор.	2-го порядка, гор /верт.	3-го порядка, гор.	3-го порядка гор /верт /диаг	3-го порядка	4-го порядка	5-го порядка	6-го порядка
A_1	0,970		0,970	0,595	0,970	0,628	0,970	0,652	0,650	0,602
A_2		0,960		0,394		0,443	0,960	0,439	0,449	0,243
A_3			0,000		0,000		0,941	—0,033	—0,032	—0,018
A_4						—0,081		—0,071	—0,065	—0,020
A_5									—0,015	—0,093
A_6										0,279
A^7					0,000					
	5,910 %	7,840 %	5,910 %	4,406 %	5,910 %	4,337 %	4,391 %	4,371 %	4,370 %	3,960 %

На рис. 22.6.2 для ряда кодирующих систем графически показано влияние изменения порядка предсказания на среднеквадратическую ошибку предсказания. Как видно, предсказание 4-го порядка уже приводит практически к минимуму этой ошибки.

В экспериментах по кодированию изображений с предсказанием было обнаружено, что статистические свойства ошибки

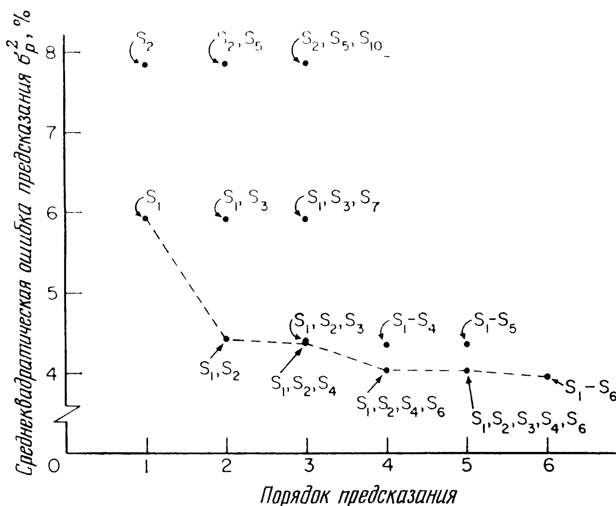


Рис. 22.6.2. Зависимость среднеквадратической ошибки предсказания в кодирующих системах с линейным предсказанием от порядка предсказания.

предсказания достаточно хорошо описываются двусторонним распределением Лапласа

$$p(D) = (\alpha/2) \exp[-\alpha |D|], \qquad (22.6.10)$$

где $\alpha = \sqrt{2}/\sigma_p$ и σ_p — среднеквадратическое отклонение разностного сигнала предсказания. Общий подход к аналитическому решению задачи о наилучшем квантовании состоит в том, чтобы, задавая число J уровней квантования, найти их величины r_j и величины $J + 1$ пороговых уровней d_j, при которых получается минимум среднеквадратической ошибки квантования. В соответствии с таким подходом расположение пороговых уровней и уровней квантования может быть рассчитано с помощью алгоритма Макса в соответствии с формулой (6.1.11).

В табл. 6.1.1 показано расположение уровней квантования и пороговых уровней для распределения Лапласа с единичной дисперсией. С другой стороны, можно воспользоваться квантова-

нием со сжатием, подвергнув разностный сигнал предсказания нелинейному преобразованию, чтобы иметь затем дело с переменной

$$z(D) = \frac{D_m [1 - \exp(-\sqrt{2/D_m}\,(D/3\sigma_p))]}{1 - \exp(-\sqrt{2D_m}/3\sigma_p)} \quad \text{при } D \geqslant 0, \tag{22.6.11а}$$

$$z(-D) = -z(D) \text{ при } D < 0. \tag{22.6.11б}$$

Величина $z(D)$ затем равномерно квантуется. Квантованная величина $z_Q(D)$ подвергается другому нелинейному преобразованию, что дает окончательно уровни квантования

$$D_Q = [\sqrt{D_m/2}\,3\sigma_p] \ln \left[1 - \frac{z_Q(D)}{D_m}\left(1 - \exp\left(-\frac{\sqrt{2D_m}}{3\sigma_p}\right)\right)\right]$$

$$\text{при } z_Q(D) \geqslant 0, \tag{22.6.12а}$$

$$D_Q = -D_Q \quad \text{при } z_Q(D) < 0. \tag{22.6.12б}$$

Множитель D_m в этих равенствах обозначает максимальный уровень квантования, определяемый алгоритмом Макса. Рассматривая этот множитель как переменный параметр, можно ценой увеличения среднеквадратической ошибки улучшить субъективную оценку качества изображений. Если установить для D_m сравнительно малый уровень, то возрастет число градаций для малых ошибок предсказания и тем самым улучшится воспроизведение тонкой структуры изображения, хотя и за счет искажений в передаче контуров с большим перепадом уровней. С другой стороны, значение D_m можно повысить, с тем чтобы улучшить передачу контуров ценой снижения чувствительности к тонким деталям изображения.

Очевидно, что экспоненциальное распределение (22.6.10) можно использовать в качестве статистической модели плотности вероятности разностного сигнала лишь тогда, когда яркость элементов изображения не приближается к верхней и нижней границам полного диапазона яркости, так что воспроизведение изображения не сопровождается ограничением яркости. Мусман [7] предложил усовершенствованную процедуру квантования, предусматривающую определение пороговых уровней и уровней квантования исходя из условной плотности вероятности $p(D \mid S_1, S_2, \ldots, S_K)$. Получаемая при этом статистическая модель позволяет учитывать, в какой степени уровень данного элемента приближается к верхней или нижней границе диапазона яркости. Построение квантователя с учетом условных плотностей вероятности высокого порядка представляется практически невыпол-

нимым из-за огромного числа возможных состояний квантуемого вектора. Однако квантователь, перестраиваемый в зависимости от близости к границам диапазона яркости, может быть осуществлен при небольшом усложнении его схемы.

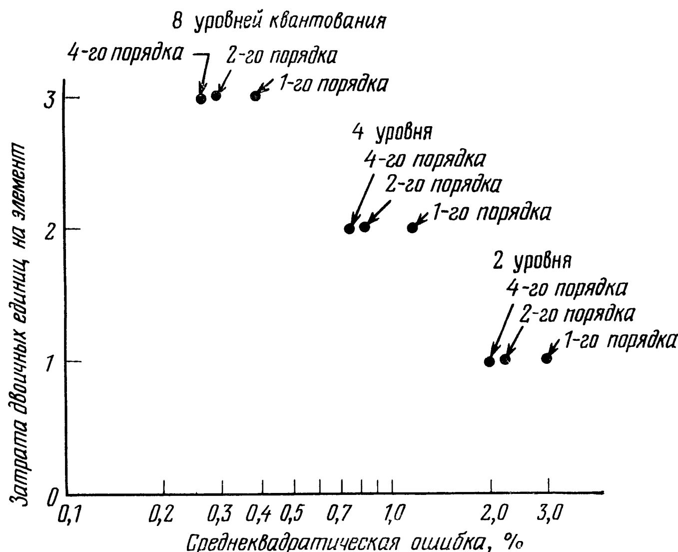


Рис. 22.6.3 Теоретическая зависимость среднеквадратической ошибки воспроизведения в кодирующих системах с линейным предсказанием от количества двоичных единиц, отводимых на передачу элемента изображения.

В соответствии с формулой (6.1.12) среднеквадратическая ошибка квантования может быть выражена через плотность вероятности ошибки предсказания следующим образом:

$$\mathscr{E} = \sigma_p^2 - \sum_{j=0}^{J-1} r_j^2 \int_{d_j}^{d_{j+1}} p(D)\,dD. \tag{22.6.13}$$

Здесь σ_p^2 — дисперсия разностного сигнала предсказания, определенная согласно (22.6.8), а d_j и r_j — пороговые уровни и уровни квантования соответственно. В случае распределения Лапласа (22.6.10) выражение для ошибки квантования имеет вид

$$\mathscr{E} = \sigma_p^2 - \sum_{j=J/2}^{J-1} r_j^2 \left[\exp(-\alpha d_j) - \exp(-\alpha d_{j+1})\right] \tag{22.6.14}$$

при условии, что число уровней квантования четно. На рис. 22.6.3 для ряда кодирующих систем с предсказанием показана зависи-

мость между расходом двоичных цифр и среднеквадратической ошибкой воспроизведения, характеризующей качество передачи изображений.

ЛИТЕРАТУРА

1. Goodall W. M., Television Transmission by Pulse Code Modulation, *Bell Syst. Tech. J.*, **30**, 33—49 (January 1951).
2. Cabrey R. L., Video Transmission over Telephone Cable Pairs by Pulse Code Modulation, *Proc. IRE*, **48**, 9, 1546—1561 (September 1960).
3. Harper L. H., PCM Picture Transmissions, *IEEE Spectrum*, **3**, 146 (June 1966).
4. Kretz F., Subjectively Optimal Quantization of Pictures, *IEEE Trans. Commun.* (Concise Paper), COM-**23**, 11, 1288—1292 (November 1975).
5. Stockham T. G., Intra-Frame Encoding for Monochrome Images by Means of a Psychophysical Model Based on Nonlinear Filtering of Multiplied Signals, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordon and Breach, New York, 1972, pp. 415—442.
6. Frei W. (частное сообщение).
7. Roberts L. G., Picture Coding Using Pseudo-Random Noise, *IRE Trans. Inf. Theory*, IT-**8**, 2, 145—154 (February 1962).
8. Lippel B., Kurland M., Marsh A. H., Ordered Dither Patterns for Coarse Quantization of Pictures, *Proc. IEEE*, **59**, 3, 429—431 (March 1971). [Имеется перевод: Липпель, Керленд, Мерш. Упорядоченный смещающий узор для грубого квантования изображений. — ТИИЭР, 1971, т. 59, № 3, с.101—102.]
9. Lippel B., Kurland M., The Effect of Dither on Luminance Quantization of Pictures, *IEEE Trans. Commun. Tech.*, COM-**19**, 6, 879—889 (December 1971).
10. Thompson J. E., Sparkes J. J., A Pseudo-Random Quantizer for Television Signals, *Proc. IEEE*, **55**, 3, 353—355 (March 1967). [Имеется перевод: Томпсон, Спаркс. Псевдослучайный квантователь телевизионного сигнала. — ТИИЭР, тематический вып. «Сокращение избыточности», 1967, т. 55, № 3, с. 122—125.]
11. Thompson J. E. A 36-MBIT/S Television Codec Employing Pseudorandom Quantization, *IEEE Trans. Commun. Tech.*, COM-19, 6, 872—879 (December 1971).
12. Limb J. O., Design of Dither Waveforms for Quantized Visual Signals, *Bell. Syst. Tech. J.*, **48**, 7, 2555—2583 (September 1969).
13. Bisignani W. T., Richards G. P., Whelan J. W., The Improved Grey Scale and the Coarse-Fine PCM Systems: Two New Digital TV Bandwidth Reduction Techniques, *Proc. IEEE*, **54**, 3, 376—390 (March 1966).
14. Richards G. P., Bisignani W. T., Redundancy Reduction Applied to Coarse-Fine Encoded Video, *Proc. IEEE*, **55**, 10, 1707—1717 (October 1967). [Имеется перевод: Ричардс, Бизиньяни. Сокращение избыточности видеосигнала, кодированного на основе грубой и тонкой шкалы. — ТИИЭР, т. 50, № 10, с. 63—75.]
15. Kretzmer E. R., Reduced-Alphabet Representation of Television Signals, *IRE Convention Record*, Part 4, 140—147 (1956).
16. Huhns M. N., Optimum Restoration of Quantized Correlated Signals, University of Southern California, Image Processing Institute, Report USCIPI 600, August 1975.
17. Huffman D. A., A Method for the Construction of Minimum-Redundancy Codes, *Proc. IRE*, **40**, 9, 1098—1101 (September 1952).
18. Davisson L. D., Kutz R. L., A Real Time Programmable Data Compression System. Proceedings International Telemetering Conference, 1971.

19. Davisson L. D., Kutz R. L., An Operational Data Compression System Using Minicomputers, National Telemetering Conference Record, Houston, Texas, December 1972, 34A-1 to 34A-4.
20. May C. L., Spencer D. J., Data Compression for Earth Resources Satellites, Proceedings International Telemetering Conference, Los Angeds, California, Vol. 8, October 1972, pp. 352—362.
21. Rice R. F., Plaunt J. R., Adaptive Variable-Length Coding for Efficien, Compression of Spacecraft Television Data, *IEEE Trans. Commun. Tech.*t **COM-19**, 6, 889—898 (December 1971).
22. Laemmel A. E., Coding Processes for Bandwidth Reduction in Picture Transmission Microwave Institute, Polytechnic Institute of Brooklyn, Report R246-251, August 1951.
23. Gouriet G. G., Bandwidth Compression of a Television Signal, *Proc. IEE*, **104**, Part B, 15, 265—272 (May 1957).
24. Michel W. S., Statistical Encoding for Text and Picture Communication, *Commun. Electron.*, **35**, 33—36 (March 1958).
25. Цуккерман И. И. О передаче координат элементов телевизионного изображения, *Радиотехника*, № 4, с. 77—89, 1958.
26. Schreiber W. F., Knapp C. F., TV Bandwidth Reduction by Digital Coding, *IRE National Convention Record*, **6**, Part 4, 88—89 (1958).
27. Schreiber W. F., Knapp C. F., Kay N. D., Synthetic Highs: An Experimental TV Bandwidth Reduction System, *J. Soc. Motion Picture and Television Engineers*, **68**, 525—537 (August 1959).
28. Julesz B., A Method of Coding TV Signals Based on Edge Detection, *Bell Syst. Tech. J.*, **38**, 4, 1001—1020 (July 1959).
29. Cherry E. C., Barton M. P., Kubba M. H., An Experimental Study Possible Bandwidth Compression of Visual Image Signals, *Proc. IEEE*, **51**, 11, 1507—1517 (November 1963). [Имеется перевод: Черри и др. Экспериментальное исследование возможного сжатия полосы частот сигналов изображения.— ТИИЭР, 1963, т. 51, № 11, с. 1479—1491.]
30. Robinson A. H., Cherry C., Results of a Prototyre Television Bandwidtя Compression Scheme, *Proc. IEEE*, **55**, 3, 356—364 (March 1967). [Имеется перевод: Робинзон, Черри. Результаты разработки опытной схемы сжатия полосы частот телевизионных сигналов. — ТИИЭР, 1967, т. 55, № 3, с. 125—135.]
31. Pratt W. K., Stop-Scan Edge Detection System for Interplanetary Televi sion Transmission, IRE National Symposium on Space Electronics and Telemetry, October 1962.
32. Capon J., A Probabilistic Model for Run-Length Coding of Pictures, *IRE Trans. Inf. Theory*, IT-5, 4, 157—163 (December 1959).
33. Pratt W. K., Coding Compression of a Television Bandwidth Reduction System, *Proc. IEEE* (*Correspondence*), **54**, 6, 914—916 (June 1966). [Имеется перевод: Прэтт. Кодовое сжатие в системе с сокращением полосы частот телевизионного сигнала. — ТИИЭР, 1966, т. 54, № 6, с. 109—110.]
34. Golomb S. W., Run Length Encodings, *IEEE Trans. Inf. Theory*, IT-**12**. 3, 399—401 (July 1966).
35. Happ W. W., Coding Schemes for Run-Length Information, Based on Poisson Distribution, National Telemetering Conference Record, April 1969, pp. 257—261.
36. Limb J. O., Efficiency of Variable Length Binary Encoding, UMR Mervin J. Kelly Communications Conference, Rolla, Missouri, October 1970, 13.3-1 to 13.3-9.
37. Beaudette C. G., An Efficient Facsimile System for Weather Graphics, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds , Gordon and Breach, New York, 1972, pp. 217—220.

38. Huang T. S., Run-Length Coding and Its Extensions, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordon and Breach, New York, 1972, pp. 221—266.
39. White H. E., Lippman M. D., Powers K. H., Dictionary Look-Up Encoding of Graphics Data, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordon and Breach, New York, 1972, pp. 265—281.
40. Schwartz J. W., Barker R. C., Bit-Plane Encoding: A Technique for Source Encoding, *IEEE Trans. Aerospace Electron. Syst.*, AES-**2**, 4, 385—392 (July 1966).
41. Spencer D. R., Huang T., Bit-Plane Encoding of Continuous-Tone Pictures, Symposium on Computer Processing in Communications, Polytechnic Institute of Brooklyn, New York, April 1969.
42. Seyler A. J., The Coding of Visual Signals to Reduce Channel-Capacity Requirements, *Proc. IEE*, **109**, C, 676—684 (September 1962).
43. Mounts F. W., A Video Encoding System Using Conditional Picture-Element Replenishment, *Bell Syst. Tech. J.*, **48**, 7, 2545—2555 (September 1969).
44. Candy J. C., Franke M. A., Haskell B. G., Mounts F. W., Transmitting Television as Clusters of Frame-to-Frame Differences, *Bell Syst. Tech. J.*, **50**, 6, 1889—1917 (July — August 1971).
45. Haskell B. G., Mounts F. W., Candy J. C., Interframe Coding of Videotelephone Pictures, *Proc. IEEE*, **60**, 7, 792—800 (July 1972).
46. Schindler H. R., Delta Modulation, *IEEE Spectrum*, **7**, 69—78 (October 1970).
47. De Jager F., Deltamodulation: A Method of PCM Transmission Using a One-Unit Code, *Philips Res. Rep.*, **7**, 442—466 (1952).
48. Salaman R. G., Digital Television Encoding, Seventh National Communication Symposium Record, 1961, pp. 274—279.
49. Balder J. C., Kramer C., Video Transmission by Delta Modulation Using Tunnel Diodes, *Proc. IRE*, **50**, 4, 428—431 (April 1962).
50. Inose H., Yasuda Y., A Unit Bit Coding Method by Negative Feedback, *Proc. IEEE*, **51**, 11, 1524—1535 (November 1963).
51. Brainard R. C., Candy J. C., Direct-Feedback Coders: Decign and Performance with Television Signals, *Proc. IEEE*, **57**, 5, 776—786 (May 1969). [Имеется перевод: Брейнард, Канди. Кодирующие устройства с прямой и обратной связью. Расчет и характеристики при работе с телевизионными сигналами.— ТИИЭР, 1969, т. 57, № 5, с. 40—52.]
52. Cutler C. C., Differential Quantization of Communication Signals, Patent 2-605-361, Application June 1950, Issuance July 1952.
53. Millard J. B., Maunsell H. I., Digital Encoding of Video Signals, *Bell Syst. Tech. J.*, **50**, 2, 459—479 (February 1971).
54. Abbott R. P., A Differential Pulse-Code-Modulation Coder for Videotelephony Using Four Bits per Sample, *IEEE Trans. Commun. Tech.*, COM-**19**, 6, 907—913 (December 1971).
55. Graham R. E., Predictive Quantizing of Television Signals, *IRE WESCON Convention Record*, Part 4, 142—157 (1958).
56. Bauch H. H. et al., Picture Coding, *IEEE Trans. Commun.*, COM-**20**, 9, 1158—1167 (September 1974).
57. Limb J. O., Source-Receiver Encoding of Television Signals, *Proc. IEEE*, **55**, 3, 364—379 (March 1967). [Имеется перевод: Лимб. Кодирование телевизионных сигналов с учетом свойств источника и получателя. — ТИИЭР, тематический вып. «Сокращение избыточности», 1967, т. 55, № 3, с. 135—152.]
58. Limb J. O., Adaptive Encoding of Picture Signals, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordon and Breach, New York, 1972, pp. 341—382.
59. Chow M. C., Variable-Length Redundancy Removal Coders for Differentially Coded Video Telephone Signals, *IEEE Trans. Commun. Tech.*, COM-**19**, 6, 922—926 (December 1971).

60. Brainard R. C., Subjective Evaluation of PCM Noice-Feedback Coder for Television, *Proc. IEEE*, **55**, 3, 346—353 (March 1967). [Имеется перевод: Брейнард. Субъективная оценка телевизионного КИМ-кодирующего устройства с шумовой обратной связью. — ТИИЭР, тематический вып. «Сокращение избыточности», 1967, т. 55, № 3, с. 114—122.]
61. Kimme E. G., Kuo F. F., Synthesis of Optimal Filters for a Feedback Quantization System, *IEEE Trans. Circuit Theory*, CT-**10**, 3, 405—413 (September 1963).
62. Harrison C. W., Experiments with Linear Prediction in Television, *Bell Syst. Tech. J.*, **31**, 4, 746—783 (July 1952).
63. O'Neal J. B., Predictive Quantizing Systems (Differential Pulse Code Modulation) for the Transmission of Television Signals, *Bell Syst. Tech. J.*, **45**, 5, 689—721 (May—June 1966).
64. Connor D. J., Pease R. F. W., Scholes W. G., Television Coding Using Two-Dimensional Spatial Prediction, *Bell. Syst. Tech. J.*, **50**, 1049—1063 (March 1971).
65. Winkler M. R., High Information Delta Modulation, *IEEE International Convention Record*, Part 8, 260—265 (1963).
66. Winkler M. R., Pictorial Transmission with HIDM, **1965** *IEEE International Convention Record*, Part 1, 285—291 (1965).
67. Bosworth R. H., Candy J. C., A Companded One-Bit Coder dor Television Transmission, *Bell Syst. Tech. J.*, **48**, 5, 1459—1479 (May—June 1969).
68. Candy J. C., Refinement of a Delta Modulator, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordon and Breach, New York, 1971, pp. 323—339.
69. Brown E. F., Sliding-Scale Operation of Differential Type PCM Codes for Television, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordon and Breach, New York, 1972, pp. 303—322.
70. Aughenbough G. W., Irwin J. D., O'Neal J. B., Delayed Differential Pulse Code Modulation, Proceedings 2nd Annual Princeton Conference, October 1970, pp. 125—130.
71. Cutler C. C., Delayed Encoding: Stabilizer for Adaptive Coders, *IEEE Trans. Commun. Tech.*, COM-**19**, 6, 898—907 (December 1971).
72. Kamunski W., Brown E. F., En Edge-Adaptive Three-Bit Ten-Level Differential PCM Coder for Television, *IEEE Trans. Commun. Tech.*, COM-**19**, 6, 944—947 (December 1971).
73. Kaul P., Golding L., A DPCM Code Using Edge Coding and Line Replacement, NTC'72 Record, IEEE Publication No. 72 CHO 601-5-NTC, Houston, Texas, December 1972, 34B-1 to 34B-6.
74. Frei A. H., Schindler H. R., Vettiger P., An Adaptive Dual-Mode Coder/Decoder for Television Signals, *IEEE Trans. Commun. Tech.*, COM-**19**, 6, 933—944 (December 1971).
75. Jayant N. S., Adaptive Delta Modulation with One-Bit Memory, *Bell Syst. Tech. J.*, **49**, 3, 321—342 (March 1970).
76. Teacher C. F., Yutz R. W., Secure Color Video Techniques, Def. Doc. Center, AD 462528, Philco Corporation, Blue Bell, Pennsylvania, February 1965.
77. Bhyshan A. K., Transmission and Coding of Color Pictures in Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordon and Breach, New York, 1972, pp. 697—725.
78. Limb J. O., Rubenstein C. B., Walsh K. A., Digital Coding of Color Picturephone Signals by Element-Differential Quantization, *IEEE Trans. Commun. Tech.*, COM-**19**, 6, 992—1006 (December 1971).
79. Habibi A., Delta Modulation and DPCM Coding of Color Signals, Proceedings of the International Telemetering Conference, Los Angeles, California, Vol. 8, October 1972, pp. 333—343.

80. Mounts F. W., Frame-to-Frame Digital Processing of TV Pictures to Remove Redundancy, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordon and Breach, New York, 1972 pp. 653—673.
81. Roese J. A., Pratt W. K., Robinson G. S., Interframe Transform Coding and Predictive Coding Methods, IEEE International Communications Conference, San Francisco, California, June 1975.
82. O'Neal J. B., Delta Modulation Quantizing Noise Analytical and Computer Simulation Results for Gaussian and TV Input Signals, *Bell Syst. Tech. J.*, **45**, 1, 117—142 (January 1966).
83. O'Neal J. B., Predictive Quantizing System (Differential Pulse Code Modulation) for the Transmission of Television Signals, *Bell Syst. Tech. J.*, **45**, 5, 689—721 (May—June 1966).
84. O'Neal J. B., A Bound on Signal-to-Quantizing Noise Ratios for Digital Encoding Systems, *Proc. IEEE*, **55**, 3, 287—292 (March 1967). [Имеется перевод: О'Нил мл. Граничное значение отношения сигнал/шум квантования в цифровых системах кодирования. — ТИИЭР, тематический вып. «Сокращение избыточности», 1967, т. 55, № 3, с. 48—54.]
85. O'Neal J. B., Entropy Coding in Speech and Television Differential PCM Systems, *IEEE Trans. Inf. Theory*, **IT-17**, 6, 758—761 (November 1971).
86. Habibi A., Comparison of Nth Order DPCM Encoder with Linear Transformation and Block Quantization Techniques, *IEEE Trans. Commun. Tech.*, **COM-19**, 6, 948—957 (December 1971).
87. Papoulis A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.
88. Bauch H. H. et al., Picture Coding, *IEEE Trans. Commun. Tech.*, **COM-20**, 9, 1158—1167 (September 1974).

Глава 23

ЦИФРОВОЕ КОДИРОВАНИЕ ИЗОБРАЖЕНИЙ С ПРОСТРАНСТВЕННОЙ ОБРАБОТКОЙ

Эта глава посвящена цифровым методам кодирования изображений, основанным на пространственной обработке. Описаны интерполяционные методы, кодирование на основе преобразования, выделение и кодирование признаков, а также символическое кодирование.

23.1. ИНТЕРПОЛЯЦИОННЫЕ МЕТОДЫ КОДИРОВАНИЯ

Интерполяционные кодирующие системы основаны на численных методах аппроксимации, посредством которых последовательность или двумерный массив отсчетов яркости приближенно представляются через непрерывные функции. Процедура интерполяции может быть применена на этапе преобразования изображения в кодированный сигнал (интерполяция на передающей стороне) или она может быть частью процесса восстановления изображения по кодированному сигналу (интерполяция на приемной стороне).

23.1.1. ИНТЕРПОЛЯЦИЯ НА ПЕРЕДАЮЩЕЙ СТОРОНЕ

В кодирующих системах с интерполяцией на передающей стороне значения яркости аппроксимируются непрерывными функциями с заранее установленной точностью [1—5]. Интерполяция может проводиться вдоль строки развертки или охватывать некоторую часть плоскости изображения.

Рис. 23.1.1, а иллюстрирует работу интерполятора нулевого порядка. В этом примере для всех элементов изображения установлен одинаковый интервал допустимых искажений, в пределах которого строится набор отрезков горизонтальных прямых максимальной длины без дополнительных ограничений на расположение их начальных и конечных точек. Всякий элемент изображения перекрывается каким-либо из этих отрезков. По каналу связи передаются вертикальная координата каждого отрезка и его длина, выраженная числом элементов. При восстановлении изображения на приемной стороне все элементы в пределах отрезка приобретают уровень, соответствующий его вертикальной координате. Этот вариант интерполяции допускает большую свободу в выборе возможных комбинаций отрезков горизонтальных прямых и

в связи с этим позволяет получить наиболее эффективное представление исходных данных с помощью минимального числа отрезков. Однако объем вычислительных операций, необходимых для построения такого оптимального приближения, часто оказы-

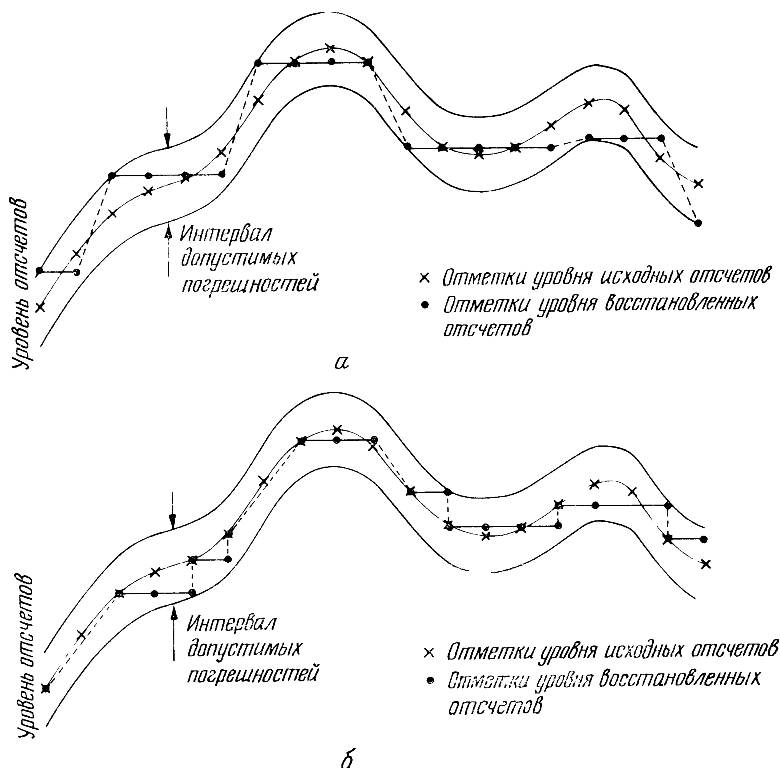


Рис. 23.1.1. Интерполяция нулевого порядка.
а — без привязки к отметкам уровня; *б* — с привязкой начальных точек к отметкам уровня.

вается слишком большим. Упрощенная схема интерполяции нулевого порядка показана на рис. 23.1.1, *б*. Здесь любой отрезок горизонтальной прямой лежит на уровне начального отсчета и оканчивается в момент взятия последнего отсчета. Упрощенный вариант интерполяции нулевого порядка представляет собой кодирование длин серий с указанием яркости первого элемента серии.

Действие различных интерполяторов первого порядка показано на рис. 23.1.2. На рис. 23.1.2, *а* всякий элемент изображения перекрывается каким-либо из прямолинейных отрезков, размещение которых в границах допустимого интервала погрешностей не связано с дополнительными ограничениями на распо-

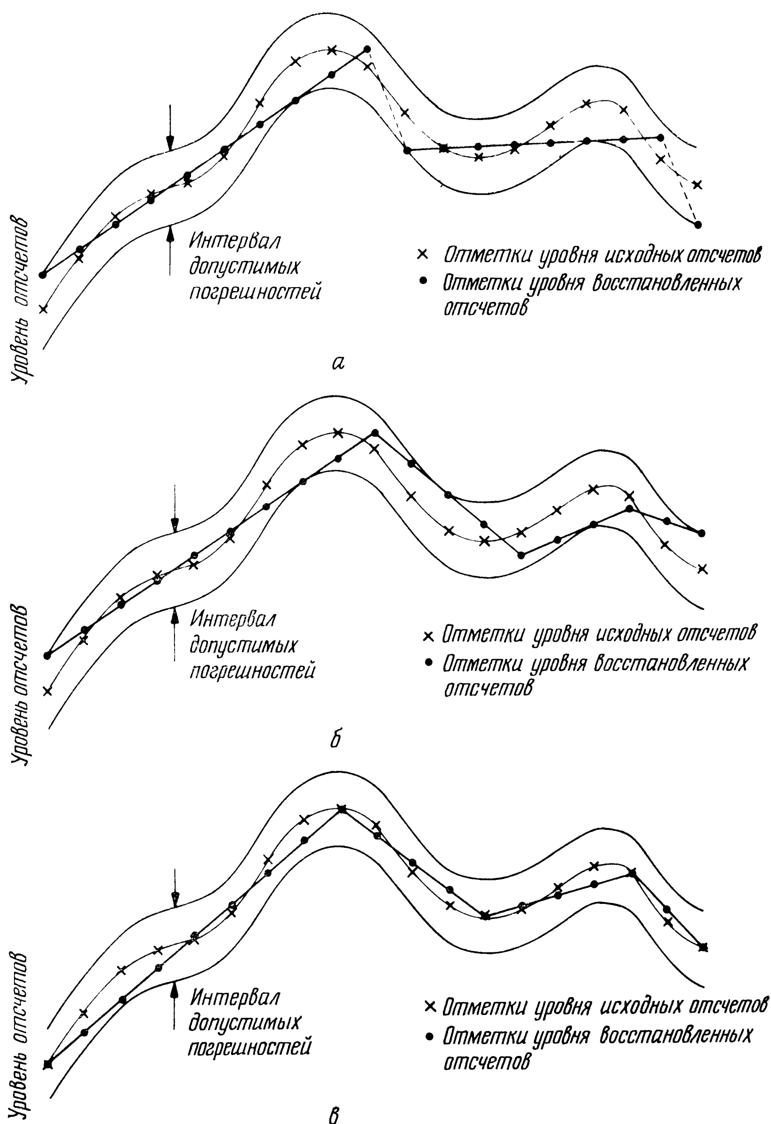


Рис. 23.1.2. Интерполяция первого порядка.

а — с независимым расположением концов каждого отрезка; *б* — с независимым расположением вершин связной ломаной относительно отметок уровня; *в* — с привязкой вершин связной ломаной к отметкам уровня.

ложение начальных и конечных точек. Вычислительная процедура аппроксимации может быть в некоторой степени упрощена соединением начала очередного отрезка с концом предшествующего отрезка (рис. 23.1.2, *б*). Дальнейшее упрощение состоит в том, чтобы в качестве начальных и конечных точек использовать значения яркости элементов, как показано на рис. 23.1.2, *в*. Приближение такого типа часто называют веерной интерполяцией.

Полиномиальные функции более высокого порядка, например кубические сплайны, также могут быть применены для кодирования с интерполяцией, однако повышение порядка полиномов сопровождается быстрым возрастанием объема вычислений. Можно также ставить задачу двумерной интерполяции нулевого и первого порядков, однако соответствующие интерполяторы трудно осуществить на практике.

Несколько очень простых алгоритмов интерполяции нулевого и первого порядка изучены теоретически [6, 7]. Анализ интерполяционных методов кодирования в общем случае с учетом ограничений наталкивается на серьезные затруднения. Экспериментальное изучение этих методов путем цифрового моделирования [2, 3] показывает, что интерполяция первого порядка дает возможность при сохранении относительно малой пиковой погрешности довести цифровой поток при передаче изображений примерно до 1,0 дв. ед./эл., но при использовании весьма сложных кодирующих систем [8].

23.1.3. ИНТЕРПОЛЯЦИЯ НА ПРИЕМНОЙ СТОРОНЕ

Полосу частот телевизионного сигнала можно сократить вдвое, если для каждого кадра передавать только нечетные строки развертки, искусственно восполняя четные строки на приемной стороне с помощью интерполяции. Можно поступить несколько иначе и передавать лишь каждый второй элемент изображения в шахматном порядке, а при восстановлении изображения применить двумерную интерполяцию. Такая система обработки телевизионных сигналов с интерполяцией на приемной стороне была разработана Габором и Хиллом [9]. Уменьшая частоту пространственной дискретизации и применяя интерполяцию в приемнике, можно получить значительное сокращение полосы частот [10]. Разреженную систему отсчетов можно сформировать просто отбрасывая определенную часть элементов изображения, например передавая только левый верхний элемент в каждой группе из 4×4 элементов. Однако если первоначальная дискретизация изображения соответствует частоте Найквиста или более низкой частоте, такая процедура приводит к искажениям за счет наложения побочных спектров. Поэтому лучше с помощью аналоговой или цифровой обработки изображения предварительно ограничить

полосу частот, а затем уже взять соответственно сниженную частоту дискретизации. Если это частота Найквиста (по отношению к полосе частот примененного фильтра), то интерполяция с применением двумерных sinc-функций или бесселевых функций (см. разд. 4.1) даст хороший результат. Тем не менее, как уже неоднократно отмечалось, эти методы интерполяции достаточно сложны для технических приложений. Более практичной альтернативой могут служить интерполяция нулевого порядка, сплайн-интерполяция и билинейная интерполяция (см. разд. 4.3).

На рис. 21.1.3 показаны примеры кодирования изображений с интерполяцией на приемной стороне, полученные посредством цифрового моделирования. На рис. 23.1.3, *б*, *в* даны результаты следующей процедуры: массив из 256×256 элементов, составлявших исходное изображение, сокращен до 64×64 элементов простым вычеркиванием 15 элементов в каждой группе размером 4×4, а для восстановления изображения применены интерполяция нулевого порядка и билинейная интерполяция. Для изображений, показанных на рис. 23.1.3, *г*, *д*, отсчеты в сокращенном массиве из 64×64 элементов были получены усреднением уровней элементов первоначального массива в пределах групп размером 4×4. Как видно, метод усреднения отсчетов в значительной мере подавляет искажения за счет наложения спектров, сопровождающие процедуру вычеркивания элементов.

23.2. КОДИРОВАНИЕ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ПРЕОБРАЗОВАНИЯ

Кодирование на основе преобразования радикально отличается от классических методов кодирования, таких, как ИКМ, кодирование с предсказанием или с интерполяцией, которые применяются непосредственно к видеосигналу. Кодирование на основе преобразования — косвенный метод. Изображение подвергается унитарному математическому преобразованию; полученные в результате коэффициенты преобразования квантуются и кодируются для передачи по каналу связи. Этот метод утвердился как эффективное и практически удобное средство кодирования одноцветных, цветных и спектрозональных изображений, в том числе и в телевизионных системах, действующих в реальном масштабе времени.

23.2.1. КОДИРОВАНИЕ ОДНОЦВЕТНЫХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ПРЕОБРАЗОВАНИЯ

Идея замены одноцветного изображения как непосредственного объекта кодирования коэффициентами его двумерного преобразования Фурье была выдвинута в 1968 г. [11, 12]. Кодирование



Рис. 23.1.3. Примеры кодирования изображений с интерполяцией на приемной стороне; сокращение цифрового потока в отношении 16 : 1.

а — оригинал, составленный из 256×256 элементов. Вычеркивание 15 элементов в каждой группе из 4×4 элементов *б* — интерполяция нулевого порядка; *в* — билинейная интерполяция. Усреднение уровней в каждой группе из 4×4 элементов: *г* — интерполяция нулевого порядка; *д* — билинейная интерполяция

посредством преобразования Фурье основано на том, что для большинства изображений естественного происхождения значения многих коэффициентов преобразования сравнительно малы. Такие коэффициенты можно часто вообще отбросить или отвести на их кодирование малое число двоичных разрядов без риска сколько-нибудь значительных искажений изображения. В 1969 г. Прэтт, Эндрюс и Кэйн обнаружили, что во многих практических случаях можно резко сократить необходимый объем вычислительных операций, если вместо преобразования Фурье воспользоваться преобразованием Адамара [13—15]. После этого были предприняты исследования по кодированию изображений с применением дискретных преобразований Карунена—Лоэва [16] и Хаара [17, 18]. Преобразование Карунена—Лоэва, известное также как преобразование Хотеллинга, обеспечивает минимальную среднеквадратическую ошибку кодирования, но требует, к сожалению, знания статистических характеристик ансамбля передаваемых изображений, и к тому же для этого преобразования нет быстрого вычислительного алгоритма. С другой стороны, преобразование Хаара характеризуется в высшей степени эффективным алгоритмом вычисления, но приводит, как правило, к сравнительно большим погрешностям кодирования. В 1971 г. Шибата и Эномото [19] предложили так называемое наклонное ортогональное преобразование векторов из 4 или 8 компонент. «Наклонный» базисный вектор, преставляющий собой дискретную пилообразную функцию, хорошо подходит для эффективного представления видеосигнала на участках строк с плавным изменением яркости. Вскоре после этого Прэтт, Чен и Уилч разработали обобщенный алгоритм наклонного преобразования векторов большой длины и двумерных массивов [20]. Как показал Ахмед, в применении к изображениям, для которых подходит марковская статистическая модель, косинусное преобразование, имеющее быстрый вычислительный алгоритм, приближается по эффективности к преобразованию Карунена — Лоэва [21]. Синусное преобразование с аналогичными свойствами предложила Джейн [22].

Все преимущества кодирования одноцветных изображений с использованием преобразований вытекают в конечном счете из особенностей распределения энергии среди коэффициентов преобразования; благодаря этим особенностям двумерный спектр изображения более удобен для кодирования, чем изображение в исходном пространственном представлении [23]. Вследствие корреляционных связей между элементами естественного изображения энергия его спектра обнаруживает тенденцию концентрироваться в относительно небольшом числе отсчетов. Рис. 23.2.1 дает наглядное представление о двумерных спектрах изображения «Портрет», полученных в результате различных унитарных преобразований. В целях сокращения полосы частот малые по

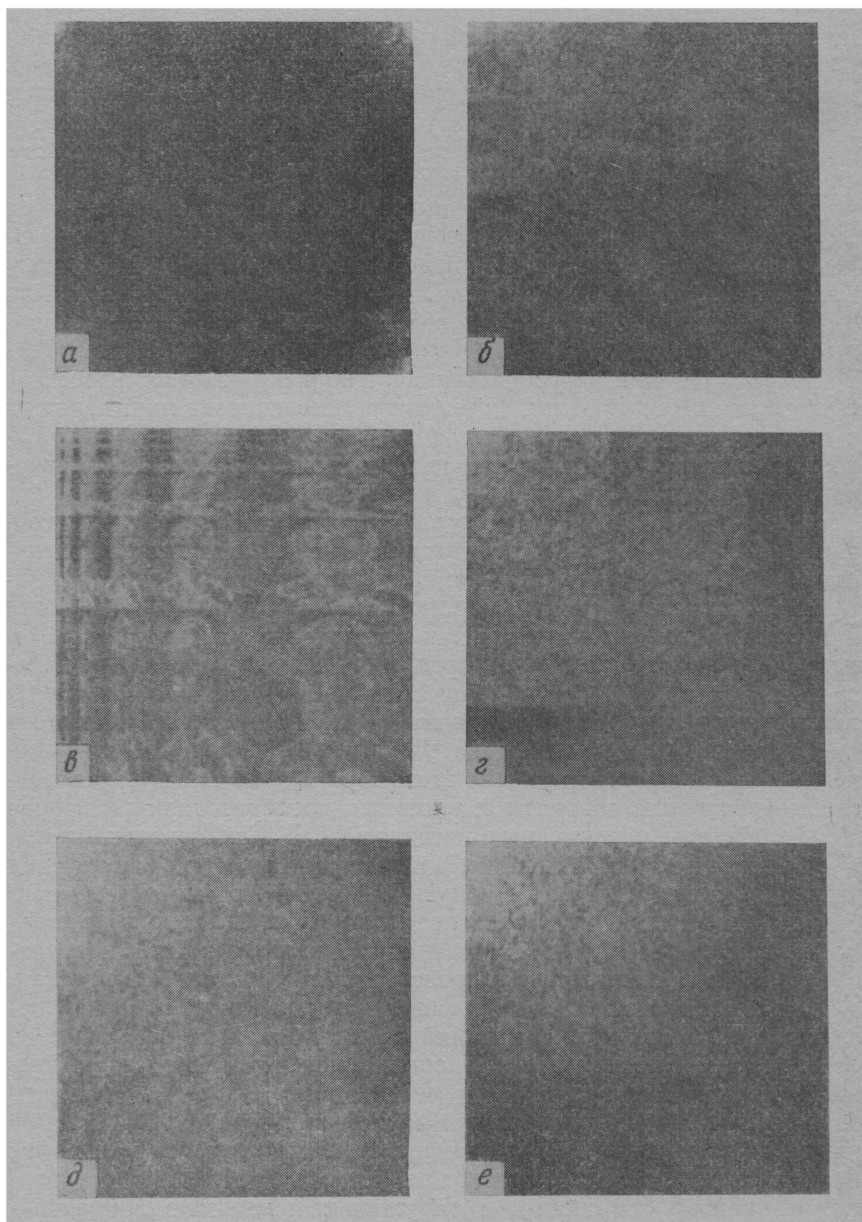


Рис. 23.2.1. Двумерные спектры изображения «Портрет», 256×256 коэффициентов. Фотоснимки воспроизводят значения логарифмов коэффициентов.

а — преобразование Фурье; *б* — преобразование Адамара; *в* — преобразование Хаара; *г* — наклонное преобразование; *д* — косинусное преобразование; *е* — синусное преобразование.

величине спектральные коэффициенты могут быть без существенного ущерба для качества изображения опущены при передаче аналоговыми средствами либо грубо проквантованы при передаче по цифровой линии связи.

Блок-схема системы кодирования посредством преобразования для передачи одноцветных изображений представлена на рис. 23.2.2. Преобразование совершается над всей совокуп-

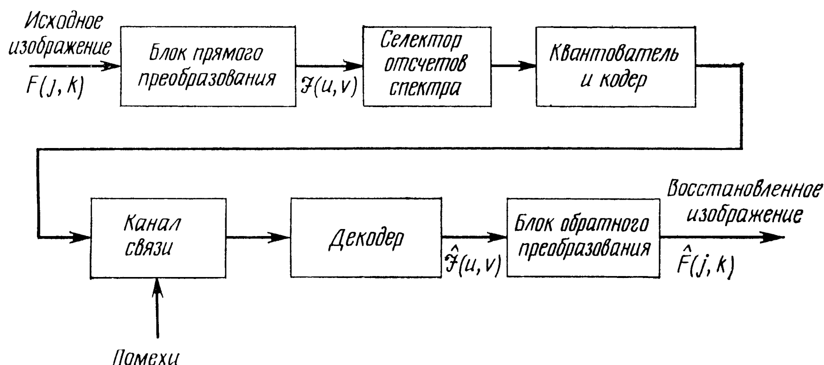


Рис. 23.2.2. Система кодирования на основе преобразования для передачи одноцветных изображений.

ностью элементов изображения или при блочном кодировании повторяется отдельно для каждого блока. Пусть числа $F(j, k)$ представляют яркости элементов блока. Коэффициенты двумерного разделимого унитарного преобразования определяются как

$$\mathcal{F}(u, v) = \sum_{j=1}^{N} \sum_{k=1}^{N} F(j, k) A_C(j, u) A_R(k, v), \qquad (23.2.1)$$

где A_R и A_C — ядра преобразований вдоль строк и столбцов. Блок элементов изображения может быть представлен матрицей $\mathbf{F}$, а результат преобразования — матрицей коэффициентов $\mathcal{F}$. Тогда преобразование записывается в матричном виде как

$$\mathcal{F} = \mathbf{A}_C \mathbf{F} \mathbf{A}_R. \qquad (23.2.2)$$

Далее производится отбор спектральных коэффициентов для передачи по каналу связи. В случае аналоговой системы связи отобранные коэффициенты перераспределяются во времени для дальнейшей равномерной передачи с помощью аналоговой модуляции. В случае цифрового канала связи отобранные отсчеты подвергаются квантованию и кодированию для передачи в виде двоичных комбинаций. На приемной стороне после декодирования поступивших данных выполняется обратное преобразование, восстанавливающее исходное изображение.

Существуют два способа отбора спектральных коэффициентов: зональный и пороговый [24]. Первый из них состоит в выделении совокупности коэффициентов, занимающих некоторые заранее очерченные, фиксированные области спектра, обычно соответствующие низкочастотным составляющим. По аналоговым линиям связи передаются значения коэффициентов, попавшие в зону отбора. При передаче по цифровым линиям связи значения отобранных коэффициентов квантуются и им присваиваются кодовые комбинации. Число уровней квантования берется пропорционально ожидаемой дисперсии соответствующего коэффициента. Второй способ отбора сохраняет лишь те коэффициенты, величина которых превышает заранее установленный порог. В связи с необходимостью сообщать при этом данные о расположении отобранных коэффициентов пороговый отбор применяется только для цифровой передачи.

Зональный отбор коэффициентов

Процесс зонального отбора в случае разделения спектра на две зоны удобно описывать с помощью селекторной функции $\mathcal{T}(u, v)$, принимающей значение 1 в области расположения передаваемых спектральных коэффициентов и значение 0 за пределами этой области. Восстановленное изображение будет теперь описываться числами

$$\hat{F}(j, k) = \sum_u \sum_v \mathcal{F}(u, v)\mathcal{T}(u, v) A_C(u, j) A_R(k, v). \quad (23.2.3)$$

Область передаваемых спектральных коэффициентов может иметь различную конфигурацию, например прямоугольную, эллиптическую или треугольную. Если следовать критерию среднеквадратической ошибки воспроизведения, то, согласно теоретическим и экспериментальным исследованиям [24], нужно придать $\mathcal{T}(u, v)$ значение 1 для тех коэффициентов, которые имеют наибольшую дисперсию.

Для теоретического анализа процесса зонального отбора представим совокупность элементов всего изображения или его блока в виде N-компонентного вектора $\mathbf{f}$, рассматриваемого как реализация случайного процесса с нулевым средним и известной ковариационной матрицей $\mathbf{K_f}$ [25]. Этот вектор подвергается линейному преобразованию, заданному матрицей $\mathbf{A}$ размера $N \times N$. N-компонентный вектор $\mathbf{f}$ обозначает результат одномерного преобразования вектора $\mathbf{f}$:

$$\mathbf{f} = \mathbf{A}\mathbf{f}. \quad (23.2.4)$$

На рис. 23.2.3 приведена схема процесса вычисления и отбора спектральных коэффициентов. Вектор $\mathscr{f}$, состоящий из всех спек-

тральных компонент, подвергается воздействию матрицы $\mathbf{S}$ размера $P \times Q$ $(P < Q)$, выделяющей определенные компоненты вектора $\mathbf{f}$. В результате получается P-компонентный вектор $\mathbf{f}_T$. Матрица $\mathbf{S}$ содержит только единицы и нули, и ее единичный элемент с номером (i, j) переводит j-ю компоненту вектора $\mathbf{f}$ в i-ю компоненту вектора $\mathbf{f}_T$. В кодирующей системе типа показанной на рис. 23.2.3 вектор $\mathbf{f}_T$ умножается затем на матрицу $\mathbf{S}^T$ размера

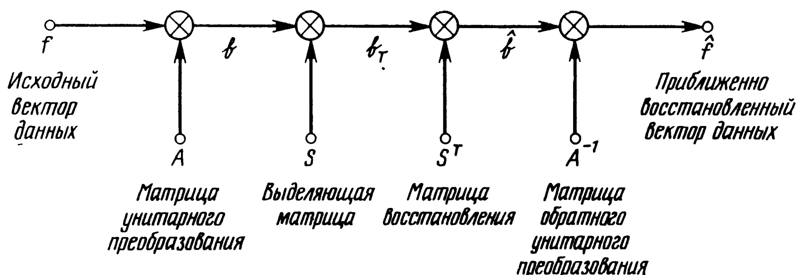


Рис. 23.2.3. Последовательность операций при зональном кодировании спектра.

$Q \times P$, в результате чего отобранным компонентам, составляющим вектор $\hat{\mathbf{f}}$, возвращаются их прежние номера, а все остальные компоненты заменяются нулями. В конечном счете операция $\mathbf{S}^T\mathbf{S}$ лишь заменяет ненужные компоненты вектора $\mathbf{f}$ нулями. Наконец, обратное унитарное преобразование $\mathbf{A}^{-1}$ дает восстановленный сигнал

$$\hat{\mathbf{f}} = \mathbf{A}^{-1}\mathbf{S}^T\mathbf{S}\mathbf{A}\mathbf{f}. \tag{23.2.5}$$

Среднеквадратическое отклонение восстановленного сигнала от исходного равно

$$\mathscr{E} = \mathbf{tr}\,[E\,\{(\mathbf{f} - \hat{\mathbf{f}})\,(\mathbf{f} - \hat{\mathbf{f}})^{*T}\}]. \tag{23.2.6}$$

Подстановка выражения (23.2.5) в (23.2.6) дает

$$\mathscr{E} = \mathrm{tr}\,[(\mathbf{I} - \mathbf{A}^{-1}\mathbf{S}^T\mathbf{S}\mathbf{A})\,\mathbf{K}_f\,[\mathbf{I} - \mathbf{A}^{-1}\mathbf{S}^T\mathbf{S}\mathbf{A}^{*T}], \tag{23.2.7а}$$

или

$$\mathscr{E} = \mathrm{tr}\,[\mathbf{A}^{-1}(\mathbf{I} - \mathbf{S}^T\mathbf{S})\,\mathbf{K}_{\mathscr{f}}\,(\mathbf{I} - \mathbf{S}^T\mathbf{S})^{*T}\,\mathbf{A}], \tag{23.2.7б}$$

если учесть, что в случае унитарных матриц связь между корреляционной матрицей $\mathbf{K}_f$ в пространственной области и корреляционной матрицей $\mathbf{K}_{\mathscr{f}}$ в спектральной области дается соотношением

$$\mathbf{K}_{\mathscr{f}} = \mathbf{A}\mathbf{K}_f\mathbf{A}^{-1}. \tag{23.2.8}$$

Так как $\operatorname{tr}\{\mathbf{XY}\} = \operatorname{tr}\{\mathbf{YX}\}$, то нетрудно показать, что

$$\mathscr{E} = \operatorname{tr}[\mathbf{K}_{\mathscr{f}}(\mathbf{I} - \mathbf{S}^T\mathbf{S})]. \tag{23.2.9}$$

Соотношение (23.2.9) выражает следующий простой результат: среднеквадратическая ошибка кодирования с зональным отбором спектральных коэффициентов равна средней энергии опущенных коэффициентов преобразования. Таким образом, среднеквадратическая ошибка в расчете на элемент изображения может быть выражена как

$$\mathscr{E} = (1/N^2) \times \sum_u \sum_v E\{\mathscr{F}^2(u, v)\} \times [1 - \mathscr{T}(u, v)], \tag{23.2.10}$$

где оператор выделения $\mathscr{T}(u, v)$ принимает значение 1 для сохраняемой компоненты спектра и значение 0 для отбрасываемой компоненты. Итак, среднеквадратическая ошибка при унитарном преобразовании определяется его способностью к концентрации энергии спектра в возможно более узких границах. В соответствии с этим, как показано в работах [26, 27], преобразование Карунена—Лоэва характеризуется наименьшей среднеквадратической ошибкой по сравнению со всеми другими унитарными преобразованиями.

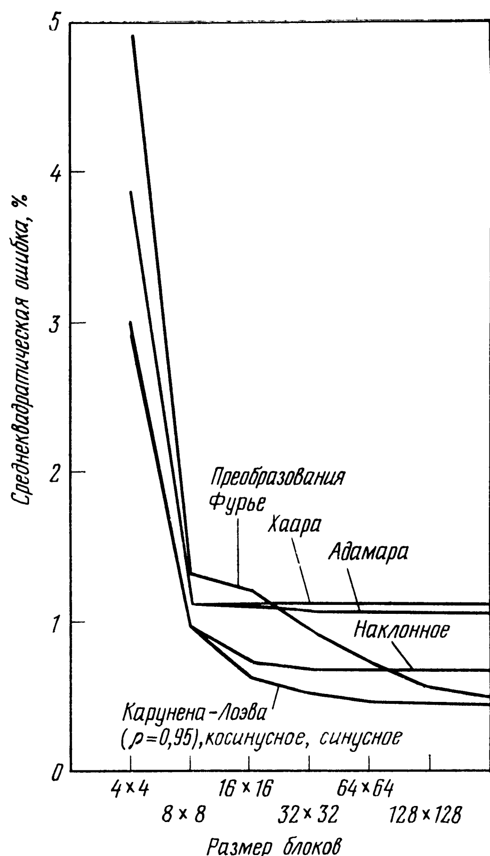


Рис. 23.2.4. Среднеквадратическая ошибка за счет зонального отбора коэффициентов как функция размера блоков.

(Зональный отбор коэффициентов с наибольшей дисперсией. Сокращение числа коэффициентов в отношении 4 · 1. $\rho_C = 0{,}95$; $\rho_R = 0{,}95$.)

На рис. 23.2.4 графически показана зависимость среднеквадратической ошибки (23.2.10) от размеров блока элементов, в пределах которого производятся преобразование и зональный



Рис. 23.2.5. Примеры кодирования с преобразованием блоков размером 16×16 элементов при зональном отборе отсчетов. Сокращение числа коэффициентов в отношении 4 : 1.

а — преобразование Фурье; *б* — преобразование Адамара; *в* — преобразование Хаара; *г* — наклонное преобразование; *д* — косинусное преобразование; *е* — преобразование Карунена—Лоэва.

отбор. Статистической моделью изображений служит при этом двумерный марковский процесс первого порядка. Отбиралось 25 % коэффициентов, имеющих наибольшую дисперсию, а остальные коэффициенты опускались. Из рис. 23.2.4 видно, что наименьшая ошибка получается в случае преобразования Карунена—Лоэва. Для марковского процесса первого порядка практически такая же минимальная ошибка получается в случае четного косинусного преобразования и синусного преобразования. Можно заметить, что увеличение блока элементов сверх размера 16×16 не приводит к дальнейшему существенному снижению среднеквадратической ошибки. Исключением является лишь преобразование Фурье, при котором с увеличением размеров блока ошибка относительно медленно приближается к минимуму, характеризующему преобразование Карунена—Лоэва. На рис. 23.2.5 приведены результаты восстановления изображения по четвертой части спектральных компонент с наибольшей дисперсией, сохраненных в результате зонального отбора после преобразований блоков размером 16×16 элементов. Субъективные оценки качества изображений, получаемых с применением различных преобразований, достаточно хорошо согласуются с оценками среднеквадратической ошибки, приведенными на рис. 23.2.4.

Зональное кодирование

В системе зонального кодирования каждый блок коэффициентов преобразования разделяется на ряд областей. Для каждой области устанавливается набор уровней квантования, число которых выбирается пропорционально средней дисперсии спектральных коэффициентов, принадлежащих данной области. Если применяется равномерный код, отводящий $N_B(u, v)$ двоичных разрядов для любого коэффициента данной области, то соответствующее число уровней квантования составит

$$L(u, v) = 2^{N_B(u, v)}. \qquad (23.2.11)$$

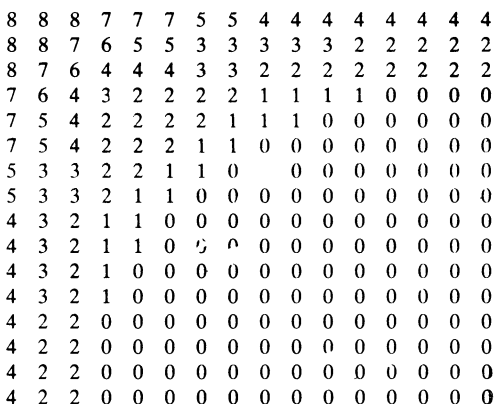
Для кодирования изображения потребуется в общей сложности

$$N_B = \sum_u \sum_v N_B(u, v) \qquad (23.2.12)$$

двоичных единиц. Число двоичных разрядов $N_B(u, v)$, отводимых на каждый коэффициент, определяется либо в соответствии с законом логарифма дисперсии (6.2.11), т. е. как

$$N_B(u, v) = (N_B/N^2) + 2\lg[V_{\mathcal{F}}(u, v)] - $$
$$- (2/N^2) \sum_{a=1}^{N} \sum_{b=1}^{N} \lg[V_{\mathcal{F}}(a, b)], \qquad (23.2.13)$$

где $V_{\mathcal{F}}(u, v)$ обозначает дисперсию коэффициента преобразования, либо с помощью последовательного алгоритма распределения двоичных разрядов, описанного в гл. 6. Если для N_B установлена малая величина, то в результате квантования некоторым коэффициентам будут приписаны нулевые значения, что равнозначно отсечению таких коэффициентов (подобно тому, как это происходит при зональном отборе). Типичное распределение двоичных разрядов между спектральными коэффициентами для блока размером 16×16 элементов показано на рис. 23.2.6.



8	8	8	7	7	7	5	5	4	4	4	4	4	4	4	4
8	8	7	6	5	5	3	3	3	3	3	2	2	2	2	2
8	7	6	4	4	4	3	3	2	2	2	2	2	2	2	2
7	6	4	3	2	2	2	2	1	1	1	1	0	0	0	0
7	5	4	2	2	2	2	1	1	1	0	0	0	0	0	0
7	5	4	2	2	2	1	1	0	0	0	0	0	0	0	0
5	3	3	2	2	1	1	0		0	0	0	0	0	0	0
5	3	3	2	1	1	0	0	0	0	0	0	0	0	0	0
4	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0
4	3	2	1	1	0	0	0	0	0	0	0	0	0	0	0
4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0
4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0

Рис. 23.2.6. Типичное распределение двоичных единиц при зональном кодировании с преобразованием блоков размером 16×16 элементов; в среднем на элемент отводится 1,5 дв. ед.

Оптимальная шкала квантования, содержащая заданное число уровней, иначе говоря, расположение пороговых уровней и уровней квантования, при котором получается минимальная среднеквадратическая ошибка, определяется алгоритмом Макса (см. гл. 6). В соответствии с формулой (6.1.11) отыскание пороговых уровней $\mathcal{D}_j(u, v)$ и уровней квантования $\mathcal{R}_j(u, v)$ сводится к совместному решению уравнений

$$\mathcal{R}_j(u, v) = 2\mathcal{D}_j(u, v) - \mathcal{R}_{j-1}(u, v), \qquad (23.2.14а)$$

$$\mathcal{R}_j(u, v) = \int\limits_{\mathcal{D}_j}^{\mathcal{D}_{j+1}} \mathcal{F}(u, v) p[\mathcal{F}(u, v)]\, d\mathcal{F} \Bigg/ \int\limits_{\mathcal{D}_j}^{\mathcal{D}_{j+1}} p[\mathcal{F}(u, v)]\, d\mathcal{F}, \qquad (23.2.14б)$$

где $p[\mathcal{F}(u, v)]$ — плотность вероятности для коэффициента, соответствующего пространственной частоте (u, v). Для коэффициента, определяющего постоянную составляющую в разложении по данному базису, в качестве статистической модели обычно используют распределение Рэлея (10.12.2). Для остальных коэффициентов преобразования достаточно точной моделью служит гауссово распределение (10.12.1). При таком подходе к кванто-

ванию и кодированию получается среднеквадратическая ошибка, соответствующая формуле (6.1.12):

$$\mathscr{E} = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \Bigg[E\{\mathscr{F}^2(u, v)\} - \\ - \sum_{j=0}^{L(u,v)} \mathscr{R}_j^2(u, v) P\{\mathscr{D}_j(u, v) < \mathscr{F}(u, v) < \mathscr{D}_{j+1}(u, v)\}\Bigg], \quad (23.2.15)$$

где $P\{\cdot\}$ обозначает вероятность того события, что значение коэффициента преобразования окажется в промежутке между j-м и соседним $(j+1)$-м пороговыми уровнями. Точное решение в замкнутой форме уравнения (23.2.15) найти не удается. На рис. 23.2.7 представлены результаты численного решения этого уравнения для ряда преобразований в случае кодирования двумерных марковских массивов с затратой в среднем 1,5 дв. ед./эл. Изображения, полученные путем цифрового моделирования процесса зонального кодирования с преобразованием блоков размером

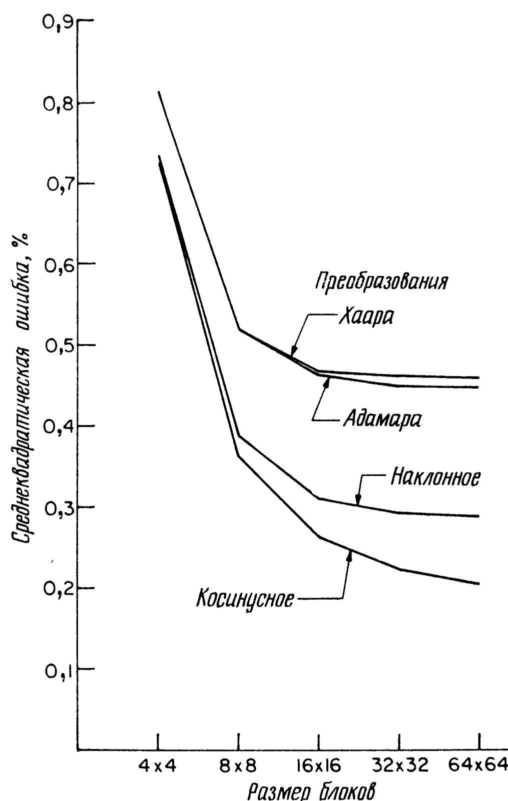


Рис. 23.2.7. Среднеквадратическая ошибка зонального кодирования в зависимости от размера блоков.

(Зональное кодирование с затратой в среднем, 1,5 дв.ед./эл.; $\rho_C = 0{,}95$, $\rho_R = 0{,}95$.)

16×16 элементов, показаны на рис. 23.2.8. Применяя к квантованным коэффициентам код Хаффмэна вместо равномерного кода, можно при той же затрате двоичных цифр несколько снизить среднеквадратическую ошибку, однако практическое построение кодера становится при этом значительно более сложной задачей.

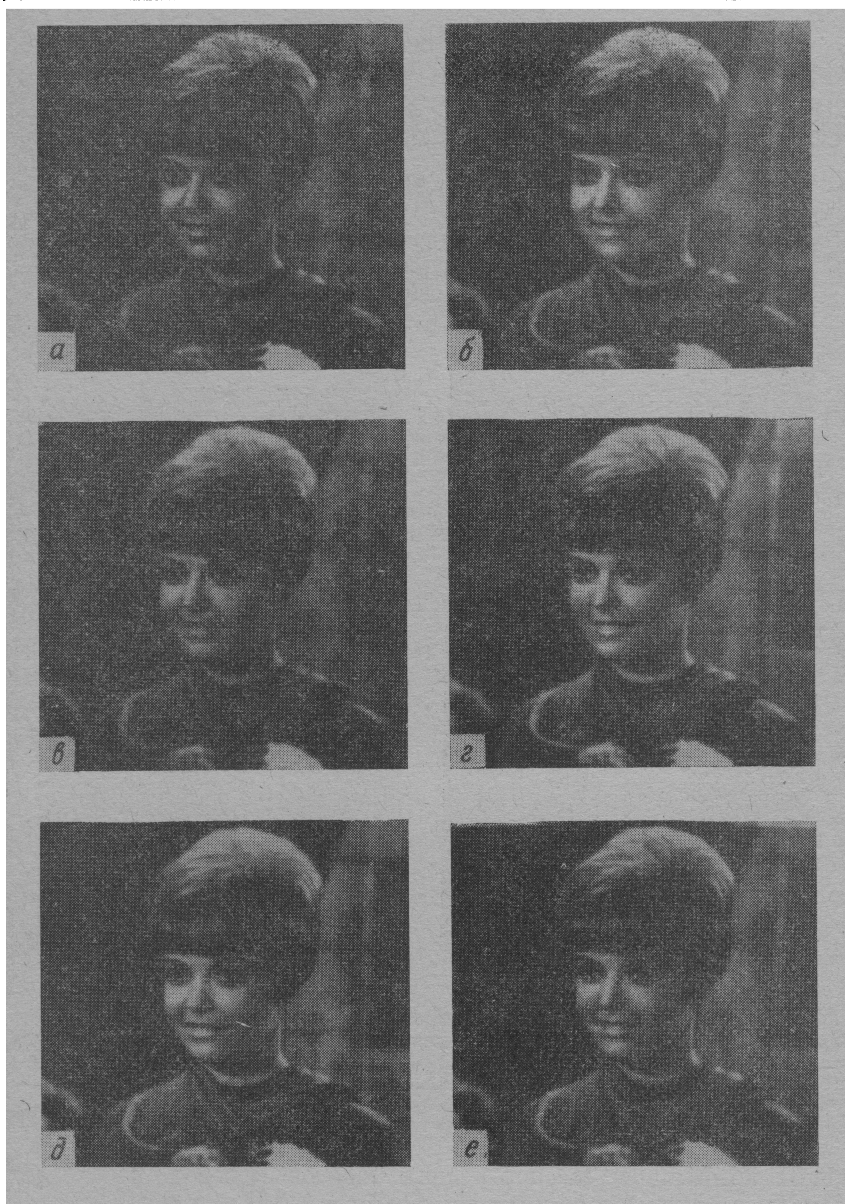


Рис. 23.2.8. Примеры зонального кодирования на основе преобразования блоков размером 16×16 элементов при средней затрате 1,5 дв. ед./эл.

а — оригинал; *б* — преобразование Адамара; *в* — преобразование Хаара; *г* — наклонное преобразование; *д* — косинусное преобразование; *е* — преобразование Карунена—Лоэва

Пороговое кодирование

В системе порогового кодирования для всех спектральных коэффициентов, превышающих по величине заданный порог, устанавливается единая шкала квантования. При передаче зна-



Рис. 23.2.9. Примеры порогового кодирования на основе преобразования Адамара и наклонного преобразования для блоков размером 16×16 элементов

а — преобразование Адамара (2,0 дв ед /эл.); *б* — преобразование Адамара (1,15 дв. ед /эл); *в* — наклонное преобразование (2,0 дв. ед /эл.); *г* — наклонное преобразование (1,15 дв. ед /эл.).

чащих коэффициентов, прошедших пороговый отбор, необходимо указывать их местоположение. Простой, но достаточно эффективный метод кодирования позиций состоит в указании числа исключенных коэффициентов, разделяющих соседние значащие коэффициенты. Общей схеме кодирования длин серий применительно к данному случаю можно придать следующий вид:

1. Первому коэффициенту каждой строки, какова бы ни была его величина, всегда ставится в соответствие кодовое слово, определенная часть которого содержит только единицы или только нули. Эта часть служит маркером для строчной синхронизации в процессе передачи.

2. Второе и последующие кодовые слова в той части, которая отведена на передачу уровня, указывают квантованную величину соответствующего значащего коэффициента. В позиционной части указывается выраженное двоичным числом количество исключенных коэффициентов, отделяющих данный значащий коэффициент от предшествующего.

3. Если в пределах заранее установленной максимальной длины серии не встретился очередной значащий коэффициент, то обе части кодового слова заполняются только единицами или только нулями, чтобы указать появление серии максимальной длины.

Введение кодовой комбинации для строчной синхронизации позволяет обходиться без кодирования номера строки, а также ограничить распространение любой ошибки, возникшей при передаче по каналу связи, пределами соответствующей строки.

Изображения, полученные путем цифрового моделирования процесса порогового кодирования, показаны на рис. 23.2.9. Адаптивный характер порогового кодирования обусловливает, как и следовало ожидать, некоторое улучшение результатов по сравнению с более простым процессом зонального кодирования Количество значащих коэффициентов, а тем самым и объем данных, передаваемых обычной системой порогового кодирования, меняются от одного изображения к другому. Для линий связи, предусматривающих передачу двоичных цифр с постоянной скоростью, разработана специальная система порогового кодирования [28]. В каждом блоке отбирается для передачи фиксированное количество N наибольших коэффициентов. Тем самым для каждого блока фактически устанавливается свой порог, обеспечивающий поддержание желаемой скорости передачи.

23.2.2. АДАПТИВНОЕ КОДИРОВАНИЕ НА ОСНОВЕ ПРЕОБРАЗОВАНИЯ

В описанных в предыдущем разделе системах кодирования одноцветных изображений шкала квантования выбирается так, чтобы минимизировать субъективно воспринимаемые искажения или среднеквадратическую ошибку воспроизведения. Однако очень часто можно существенно повысить эффективность кодирования на основе преобразований, перестраивая шкалу квантования в соответствии с локальными характеристиками различных частей передаваемого изображения.

Одно из важнейших свойств преобразования как средства кодирования заключается в том, что оно позволяет представить пространственное сильно коррелированное изображение в виде последовательности почти некоррелированных коэффициентов. Может показаться, что если такая цель достигнута, то нет смысла использовать при квантовании некоторого коэффициента данные о соседних коэффициентах. Следует напомнить, однако, что коэффициенты преобразования, даже если они почти полностью лишены корреляционных связей, сохраняют взаимную статистическую зависимость. Это обстоятельство может, таким образом, служить основанием для адаптивных процедур квантования и восстановления коэффициентов преобразования.

Адаптивный метод квантования в случае преобразования Фурье, предусматривающий раздельное адаптивное кодирование модуля и фазового сдвига каждого коэффициента, разработан Эндрюсом и Тешером [29]. При этом предполагалось, что фазовый сдвиг $\Phi(u, v)$ коэффициентов имеет равномерное распределение, а модуль $\mathcal{M}(u, v)$ — распределение Рэлея. Число уровней квантования модуля и фазового сдвига устанавливалось пропорционально дисперсии $V_{\mathcal{F}}(u, v)$ с дополнительным условием, что число уровней квантования фазового сдвига вдвое больше числа уровней квантования модуля. Среднее значение модуля коэффициента пропорционально его дисперсии:

$$E\{\mathcal{M}(u, v)\} = \overline{\mathcal{M}}(u, v) = (\pi/2)\, V_{\mathcal{F}}(u, v), \qquad (23.2.16)$$

поэтому выборочная оценка модулей коэффициентов может использоваться для оценки их дисперсий. Оценка модуля каждого коэффициента подсчитывалась как среднее квантованных модулей четырех соседних коэффициентов по формуле

$$\overline{\mathcal{M}}(u, v) = {}^1/_4((\widehat{\mathcal{M}}(u, v-1) + \widehat{\mathcal{M}}(u-1, v-1) + \widehat{\mathcal{M}}(u-1, v) + \widehat{\mathcal{M}}(u-1, v+1)). \qquad (23.2.17)$$

Этот метод адаптивного кодирования коэффициентов был распространен также на преобразование Адамара. Результаты цифрового моделирования указывают на снижение среднеквадратической ошибки и улучшение субъективной оценки качества изображений по сравнению с кодированием без адаптации.

Тасто и Уинц [30, 31] исследовали систему адаптивного кодирования на основе преобразования с разделением блоков изображения на три категории в зависимости от количества деталей и средней яркости. Для блоков каждой категории использовалась своя статистическая модель. Первую категорию составляли все блоки со значительным числом деталей, вторую и третью — блоки с малым числом деталей, соответственно более светлые и более

темные, чем средний блок этих категорий. Подобные методы были разработаны Ченом [32], а также Коксом и Тешером [33]. Моделирование адаптивных систем такого типа показало, что с их помощью можно дополнительно сократить объем передаваемых данных вдвое по сравнению с неадаптивными кодерами лучших конструкций; однако адаптивные кодеры, вообще говоря, весьма сложны в практическом осуществлении.

23.2.3. КОДИРОВАНИЕ ЦВЕТНЫХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ПРЕОБРАЗОВАНИЯ

Блок-схема системы кодирования на основе преобразования для передачи цветных изображений приведена на рис. 23.2.10 [34]. В этой системе цветное изображение представлено сигналами координат цвета источника $R(j, k)$, $G(j, k)$, $B(j, k)$, определяющими красную, зеленую и синюю составляющие для каждого элемента изображения с координатами (j, k) в системе координат приемника НТСЦ. Затем система координат цвета источника

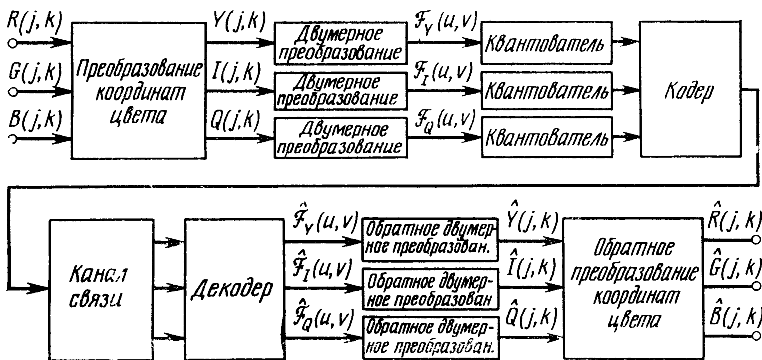


Рис. 23.2.10. Система кодирования на основе преобразования для передачи цветных изображений.

преобразуется в другую трехмерную систему координат $Y(j, k)$, $I(j, k)$, $Q(j, k)$, определяющую сигналы яркости и цветности элемента изображения, которые передаются в системе НТСЦ. Преобразование координат определяется следующим образом:

$$\begin{bmatrix} Y(j,k) \\ I(j,k) \\ Q(j,k) \end{bmatrix} = \begin{bmatrix} 0{,}299 & 0{,}587 & 0{,}114 \\ 0{,}596 & -0{,}274 & -0{,}322 \\ 0{,}211 & -0{,}523 & 0{,}312 \end{bmatrix} \begin{bmatrix} R(j,k) \\ G(j,k) \\ B(j,k) \end{bmatrix}. \qquad (23.2.18)$$

Сигналы Y, I, Q удобнее кодировать, чем координаты цвета R, G, B, по той причине, что эти сигналы в значительной степени

декоррелированы и основная часть их энергии приходится на компоненту Y. В результате появляется возможность спроектировать более эффективный квантователь. Табл. 23.2.1 позволяет сравнить распределения энергии сигналов в системах координат RGB и YIQ.

Таблица 23.2.1

Сравнение преобразований цветовых координат по степени концентрации энергии

Испытательное изображение	Система координат	σ_1^2, %	σ_2^2, %	σ_3^2, %
Портрет	RGB	45,14	35,41	19,45
	YIQ	78,32	17,54	4,14
	$K_1K_2K_3$	85,84	12,10	2,06
Два лица	RGB	51,55	31,09	17,36
	YIQ	84,84	13,81	1,35
	$K_1K_2K_3$	92,75	6,46	0,79

Каждая из компонент Y, I и Q в отдельности подвергается затем двумерному преобразованию для кодирования, что дает три двумерных спектра $\mathcal{Y}(u, v)$, $\mathcal{I}(u, v)$ и $\mathcal{Q}(u, v)$:

$$\mathcal{Y} = \mathbf{A}\mathbf{Y}\mathbf{A}^T, \tag{23.2.19а}$$

$$\mathcal{I} = \mathbf{A}\mathbf{I}\mathbf{A}^T, \tag{23.2.19б}$$

$$\mathcal{Q} = \mathbf{A}\mathbf{Q}\mathbf{A}^T. \tag{23.2.19в}$$

Здесь $\mathbf{A}$ — матрица преобразования. Следует отметить, что поскольку преобразование координат цвета и двумерное преобразование для кодирования являются линейными операциями, очередность их проведения может быть изменена. Далее производится квантование спектральных коэффициентов, причем число уровней квантования устанавливается пропорциональным ожидаемой дисперсии соответствующего коэффициента; расположение уровней также подбирается индивидуально для минимизации среднеквадратической ошибки квантования. Квантованные спектральные коэффициенты $\hat{\mathcal{Y}}(u, v)$, $\hat{\mathcal{I}}(u, v)$, $\hat{\mathcal{Q}}(u, v)$ кодируются и передаются по каналу связи. На приемной стороне производится декодирование, после чего выполняются обратные двумерные преобразования

$$\hat{\mathbf{Y}} = \mathbf{A}^T\widehat{\mathcal{Y}}\mathbf{A}, \tag{23.2.20а}$$

$$\hat{\mathbf{I}} = \mathbf{A}^T\widehat{\mathcal{I}}\mathbf{A}, \tag{23.2.20б}$$

$$\hat{\mathbf{Q}} = \mathbf{A}^T\widehat{\mathcal{Q}}\mathbf{A}. \tag{23.2.20в}$$

Полученные результаты подвергаются обратному трехмерному преобразованию, что дает восстановленные значения координат цвета:

$$\begin{bmatrix} \hat{R}(j, k) \\ \hat{G}(j, k) \\ \hat{B}(j, k) \end{bmatrix} = \begin{bmatrix} 1{,}000 & 0{,}956 & 0{,}621 \\ 1{,}000 & -0{,}272 & -0{,}647 \\ 1{,}000 & -1{,}106 & 1{,}703 \end{bmatrix} \begin{bmatrix} \hat{Y}(j, k) \\ \hat{I}(j, k) \\ \hat{Q}(j, k) \end{bmatrix}. \qquad (23.2.21)$$

Эффективность системы в целом зависит от того, насколько хорошо выполняет свою роль каждое из преобразований, производимых в процессе кодирования. Трехмерное преобразование координат цвета перераспределяет энергию между тремя компонентами колориметрического описания изображения с целью ее концентрации. Двумерное преобразование перераспределяет затем энергию каждой компоненты в отдельности, концентрируя ее в узкой области спектра. В соответствии с этой точкой зрения оптимальным трехмерным преобразованием будет преобразование Карунена—Лоэва, обеспечивающее полную декорреляцию $3N^2$ элементов цветного изображения. Показано [34], что преобразование координат цвета к системе Y, I, Q почти не уступает по степени концентрации энергии оптимальному преобразованию, т. е. преобразованию Карунена—Лоэва. Этот результат подтверждается приведенными в табл. 23.2.1 данными о распределении энергии для двух испытательных цветных изображений.

Для того чтобы оптимальным образом спроектировать систему кодирования цветных изображений на основе преобразования, необходимо исходить из аналитического критерия верности воспроизведения цветного изображения. Но установившихся критериев такого рода, к сожалению, нет. Выход из положения находят, как правило, в том, чтобы спроектировать систему квантования коэффициентов преобразования с минимальным среднеквадратическим отклонением восстановленных координат цвета $\hat{Y}$, $\hat{I}$, $\hat{Q}$ от переданных координат цвета Y, I, Q. Это отклонение определяется как

$$\mathscr{E} = (1/3N^2) \sum_j \sum_k \{[Y(j, k) - \hat{Y}(j, k)]^2 + [I(j, k) - \hat{I}(j, k)]^2 + [Q(j, k) - \hat{Q}(j, k)]^2\}. \qquad (23.2.22)$$

За последнее время изучено несколько вариантов процедуры квантования. Как выяснилось, в высшей степени эффективна по своим результатам следующая последовательность действий:

1. В качестве статистической модели для матриц ковариации координат цвета R, G, B вдоль строк и столбцов берут марков-

скую последовательность первого порядка и производят численную оценку дисперсий сигналов Y, I, Q.

2. Постоянная составляющая спектра $\mathcal{Y}$ описывается распределением Рэлея; другие составляющие спектра $\mathcal{Y}$, а также составляющие спектров $\mathcal{I}$ и $\mathcal{Q}$ — нормальными распределениями с дисперсиями, полученными в п. 1.

3. Число уровней квантования для каждой из компонент спектров Y, I, Q принимают равным

$$L_Y(u, v) = 2^{N_Y(u, v)}, \quad (23.2.23а)$$

$$L_I(u, v) = 2^{N_I(u, v)}, \quad (23.2.23б)$$

$$L_Q(u, v) = 2^{N_Q(u, v)}, \quad (23.2.23в)$$

причем количество двоичных разрядов, отводимых на кодирование каждой из компонент, устанавливают пропорционально логарифму ее дисперсии, подсчитанной в п. 1.

4. Полное число двоичных разрядов, отводимых на кодирование цветного изображения, устанавливается равным

$$N_B = N_{BY} + N_{BI} + N_{BQ}, \quad (23.2.24)$$

где

$$N_{BY} = \sum_{u=1}^{N} \sum_{v=1}^{N} N_Y(u, v), \quad (23.2.25а)$$

$$N_{BI} = \sum_{u=1}^{N} \sum_{v=1}^{N} N_I(u, v), \quad (23.2.25б)$$

$$N_{BQ} = \sum_{u=1}^{N} \sum_{v=1}^{N} N_Q(u, v). \quad (23.2.25в)$$

5. Для данного N_B выбирают три величины N_{BY}, N_{BI}, N_{BQ} и вычисляют количество уровней квантования в соответствии с соотношениями п. 3, а затем производят квантование каждого коэффициента с использованием неравномерной шкалы, составленной при помощи алгоритма Макса на основе статистических моделей п 2.

6. Вычисляют среднеквадратическую ошибку воспроизведения испытательных цветных изображений, а затем, пользуясь итерационными методами поиска минимума, определяют оптимальное распределение общего числа отведенных на кодирование двоичных разрядов между тремя компонентами.

Следует отметить, что описанную процедуру нет необходимости повторять отдельно для каждого передаваемого изображения. Оптимизация должна быть проведена лишь один раз для типичных цветных изображений соответствующей категории, с тем чтобы

получить исходные данные для проектирования конкретной схемы квантователя.

На рис. 23.2.11 показана зависимость среднеквадратической ошибки воспроизведения цветного изображения «Портрет» от

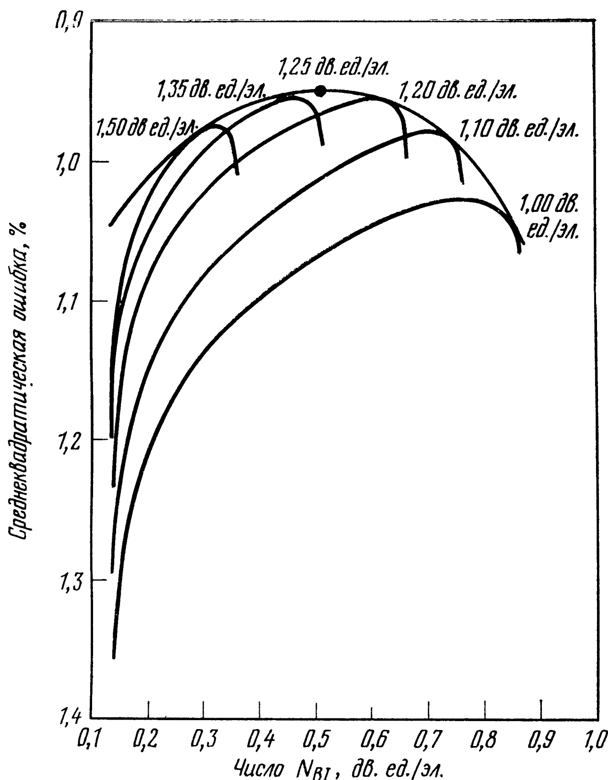


Рис. 23.2.11. Среднеквадратическая ошибка квантования цветного изображения при различных вариантах распределения двоичных единиц между его компонентами ($N_{BY} + N_{BI} + N_{BQ} = 2$ дв. ед /эл.).

величины N_{BI} при различных N_{BY} и фиксированном значении N_B. Как выяснилось, для данного изображения оптимальным является следующее распределение среднего числа двоичных разрядов между компонентами: $N_{BY} = 1{,}25$, $N_{BI} = 0{,}55$, $N_{BQ} = 0{,}20$. Это распределение практически не меняется при переходе к другим изображениям или при изменении общего числа двоичных разрядов, отводимых на передачу.

На рис. 23.2.12 представлены результаты цифрового моделирования системы кодирования с применением наклонного преоб-

а


R


G


B

б


Y


I


Q

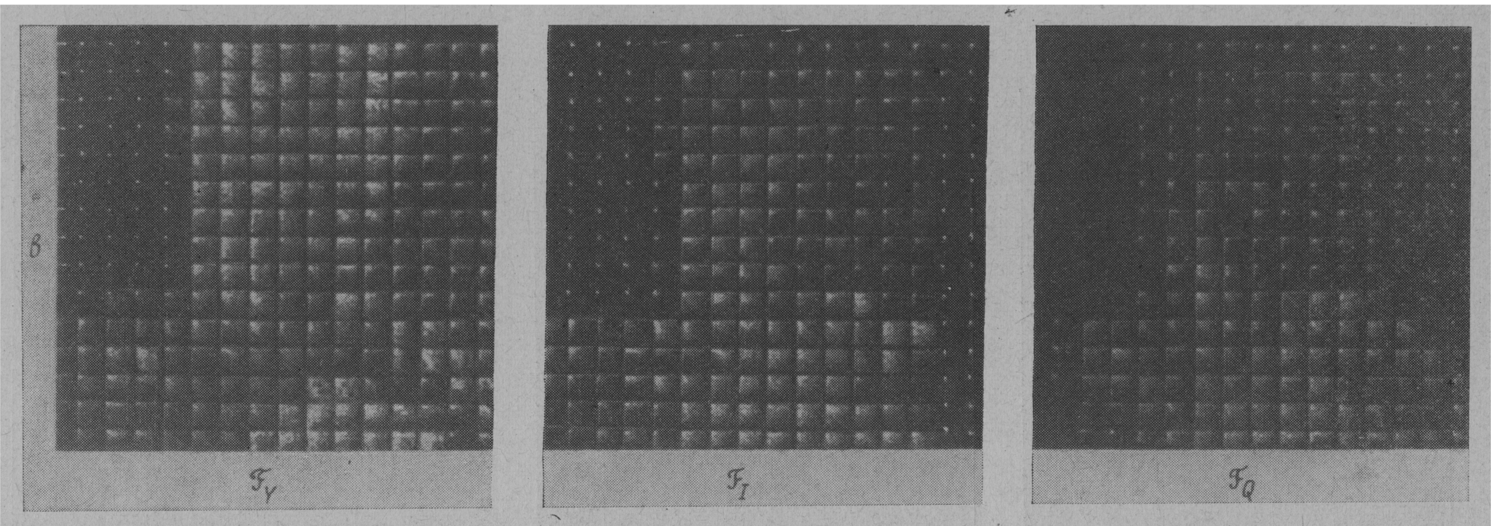
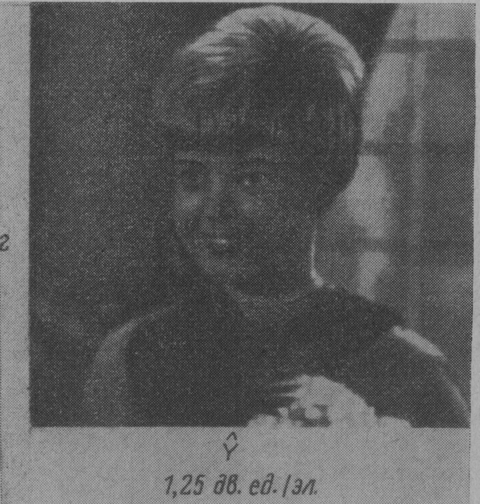


Рис. 23.2.12. Пример кодирования цветного изображения на основе наклонного преобразования при затрате в среднем 2,0 дв. ед. на элемент.

г



$\hat{Y}$

1,25 дв. ед./эл.



$\hat{I}$

0,55 дв. ед/эл.



$\hat{Q}$

0,20 дв. ед./эл.



Рис. 23.2.12 (*Продолжение*)

разования, затрачивающей на передачу элемента цветного изображения в среднем 2,0 дв. ед. Компоненты R, G, B и Y, I, Q цветного изображения на рис. 23.2.12*а*, *б* показаны в виде одноцветных фотоснимков. Следует отметить, что даже визуально ощущается высокий уровень корреляции между компонентами R, G, B. Фотоснимки на рис. 23.2.12, *в* воспроизводят значения логарифмов коэффициентов наклонного преобразования, примененного к блокам размером 16×16 элементов, и иллюстрируют эффект концентрации энергии в спектрах каждой из компонент Y, I, Q. Результаты воспроизведения компонент Y, I, Q и R, G, B после зонального кодирования показаны соответственно на рис. 23.2.12, *г* и *д*. Как видно, кодирование компонент Y, I, Q приводит к определенным искажениям, которые, однако, при воспроизведении цветного изображения не сказываются существенно на визуальной оценке изображения вследствие ограниченной полосы пространственных частот, пропускаемых зрительной системой человека.

Эномото и Шибата [35] использовали другой подход к кодированию цветного изображения, состоящий в применении преобразования в полному цветовому телевизионному сигналу, яркостная и цветностные компоненты которого объединяются методом частотного перемежения. Преобразование Фурье полного сигнала дает спектр, энергия которого концентрируется вокруг частоты поднесущей сигнала цветности. Как следствие высокочастотные составляющие спектра приобретают тенденцию к увеличению и требуют поэтому большего числа уровней квантования, чем в одноцветной системе передачи. Хорошее качество передачи цветного изображения достигается в этом случае при затрате в среднем от 3,0 до 4,0 дв. ед./эл.

23.2.4. МЕЖКАДРОВОЕ КОДИРОВАНИЕ НА ОСНОВЕ ПРЕОБРАЗОВАНИЯ

Большинство изображений, передаваемых телевизионными системами, характеризуется высоким уровнем как межкадровой (временно́й), так и внутрикадровой (пространственной) корреляции. В связи с этим можно использовать кодирование на основе трехмерных преобразований, обеспечивающих эффекты декорреляции и концентрации энергии по отношению как к отдельному кадру, так и к последовательности кадров.

Пусть $F(j, k, i)$ представляет уровни блока из $J\times K\times I$ элементов, выделенного в последовательности I кадров. Разделимое трехмерное унитарное преобразование такого блока определяется следующим образом:

$$\mathcal{F}(u, v, w)=\sum_{j=1}^{I}\sum_{k=1}^{K}\sum_{i=1}^{I} F(j, k, i)\, A_C(j, u)\, A_R(k, v)\, A_T(i, w). \qquad (23.2.26)$$

Здесь $A_C(j, u)$, $A_R(k, v)$ и $A_T(i, w)$ — ядра преобразований, производимых соответственно по столбцам, строкам и вдоль временной координаты. При построении кодирующих систем с применением трехмерных преобразований можно использовать те же

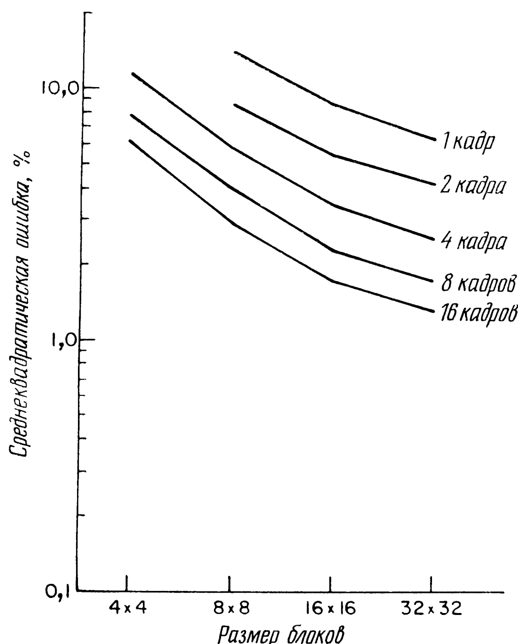


Рис. 23.2.13. Среднеквадратическая ошибка при зональном отборе коэффициентов как функция размера блоков при межкадровом кодировании на основе косинусного преобразования.

(Зональный отбор коэффициентов. Сокращение числа коэффициентов в отношении 32 : 1. Коэффициент корреляции вдоль пространственных координат $\rho_S = 0{,}96$. Коэффициент корреляции вдоль временно́й координаты $\rho_T = 0{,}96$.)

методы зонального отбора коэффициентов, зонального кодирования или порогового кодирования, какие применяются в случае двумерных преобразований. Массив $\hat{\mathcal{F}}(u, v, w)$ коэффициентов трехмерного преобразования, восстановленный на приемной стороне, подвергается обратному преобразованию, что дает уровни $\hat{F}(j, k, i)$ блока восстановленных элементов.

В процессе межкадрового кодирования для операций вдоль временно́й координаты может быть применено любое унитарное преобразование; необходимо только, чтобы длина последовательности кадров была согласована с данным преобразованием, например выбиралась бы среди чисел $I = 2, 4, 8, \dots$ Одно из свойств

преобразования Фурье может оказаться особенно полезным для межкадрового кодирования, а именно свойство реагировать на сдвиг объекта преобразования одним лишь фазовым сдвигом коэффициентов. Если последовательные кадры отличаются от начального только смещением некоторого объекта в поле зрения,

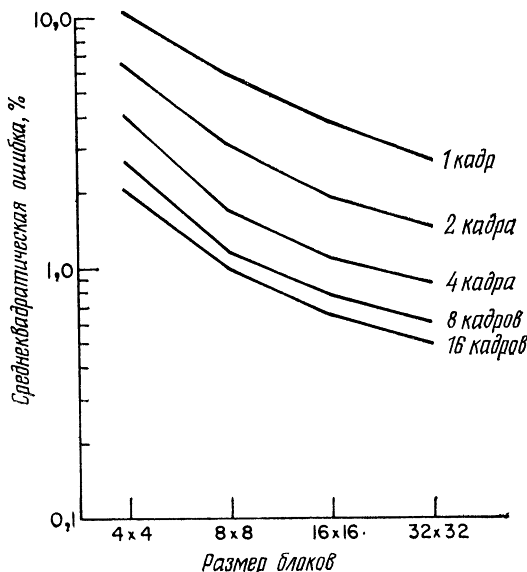


Рис. 23.2.14. Среднеквадратическая ошибка за счет зонального кодирования как функция размера блоков при межкадровом кодировании на основе косинусного преобразования.

(Зональное кодирование при средней затрате 0,25 дв. ед. на элемент. $\rho_S = 0{,}96$, $\rho_T = 0{,}96$.)

то теоретически возможно, указав амплитуду и фазовый угол коэффициентов для начального кадра, ограничиться далее лишь относительным фазовым сдвигом коэффициентов. Что касается среднеквадратической ошибки межкадрового кодирования, то различные унитарные преобразования связаны между собой точно так же, как в случае двумерного кодирования. На рис. 23.2.13 и 23.2.14 для трехмерного марковского источника сигналов показана теоретическая зависимость среднеквадратической ошибки межкадрового кодирования от размера блоков для случаев применения зонального отбора отсчетов и зонального кодирования [36].

Возможность использования преобразований Фурье и Адамара для межкадрового кодирования изучалась Эндрюсом и Тешером [29]. Принятая ими процедура кодирования заключалась в дву-

мерном преобразовании каждого кадра в целом и применении затем преобразования вдоль временно́й координаты к пространственным спектрам, соответствующим четырем последовательным



Рис. 23.2.15. Примеры межкадрового кодирования на основе косинусного преобразования. Последний, 16-й по счету кадр блока. НСКО — нормированная среднеквадратическая ошибка.

а) 1,0 дв. ед /эл , НСКО = 0,16 %; *б*) 0,5 дв ед /эл., НСКО = 0,30 %; *в*) 0,25 дв. ед /эл., НСКО = 0,49 %; *г*) 0,1 дв. ед /эл., НСКО = 0,88 %.

кадрам. Как показало цифровое моделирование, в случае видеотелефонных изображений преобразование Фурье обеспечивает возможность кодирования со среднеквадратической ошибкой около 2 % при затрате в среднем 0,25 дв. ед. на элемент или 1 % при затрате 0,5 дв. ед. на элемент. Преобразование Адамара в этом случае не позволяет достигнуть столь же хороших результатов.

Роуз и др. [37] изучали эффективность применения трехмерного преобразования Фурье и косинусного преобразования к кубическим блокам размером 16×16×16 элементов для кодирования изображений движущихся объектов, передаваемых подвижной камерой. Как можно судить по изображениям на рис. 23.2.15, отличные результаты достигались при снижении цифрового потока вплоть до 0,25 дв. ед./эл.

23.3. УМЕНЬШЕНИЕ ОШИБОК КВАНТОВАНИЯ ПРИ КОДИРОВАНИИ НА ОСНОВЕ ПРЕОБРАЗОВАНИЯ

Один из основных приемов, используемых при кодировании изображений на основе преобразования, состоит в отборе и передаче только тех коэффициентов преобразования, которые занимают заранее определенную зону двумерного спектра (обычно это низкочастотные компоненты). Перед тем как выполнить обратное преобразование на приемной стороне, в таких случаях обычно заменяют все опущенные при передаче коэффициенты нулями. Основываясь, однако, на данных о корреляционных связях в принятой совокупности коэффициентов, можно приближенно восстановить отсеченные коэффициенты вместо того, чтобы произвольно приравнивать их нулю. Процесс такого восстановления называют экстраполяцией спектра [38]. Другой алгоритм кодирования на основе преобразования предусматривает для любого коэффициента определенную шкалу квантования, число уровней которой пропорционально ожидаемой дисперсии коэффициента. Этот алгоритм кодирования вносит, разумеется, искажения. Применяя кодирование на основе преобразования, шум квантования, как правило, принимают таким, каков он есть. Тем не менее, поскольку между коэффициентами существуют корреляционные связи, можно снизить шум квантования, прибегнув к процедуре типа интерполяции спектра [39].

Экстраполяция спектра

На рис. 23.3.1 показана диаграмма, поясняющая процесс зонального кодирования с восстановлением изображения посредством экстраполяции спектра. В результате зонального кодирования вектор $\mathbf{f}$, соответствующий исходному изображению, преобразуется с помощью унитарной матрицы $\mathbf{A}$ в вектор $\mathscr{f}$, составляющими которого являются коэффициенты преобразования. Умножение полученного вектора $\mathscr{f}$ на выделяющую матрицу $\mathbf{S}$ размера $P \times Q$ $(P < Q)$ дает усеченный вектор $\mathscr{f}_T$ коэффициентов размера $P \times 1$. Декодер производит умножение вектора $\mathscr{f}_T$ на матрицу $\mathbf{W}$, в результате чего происходит приближенное восстановление отсеченных коэффициентов. Затем с помощью обратного

унитарного преобразования $\mathbf{A}^{-1}$ получается оценка $\hat{\mathbf{f}}$ исходного вектора $\mathbf{f}$. Восстанавливающая матрица $\mathbf{W}$ выбирается так, чтобы минимизировать среднеквадратическое отклонение вектора $\mathbf{f}$ от вектора $\hat{\mathbf{f}}$. Для отыскания оптимальной матрицы $\mathbf{W}$ следует потребовать, чтобы в среднем для различных изображений отклонение ($\mathbf{f} - \hat{\mathbf{f}}$) было ортогонально усеченному вектору $\mathscr{f}_T$. Таким образом, полагая

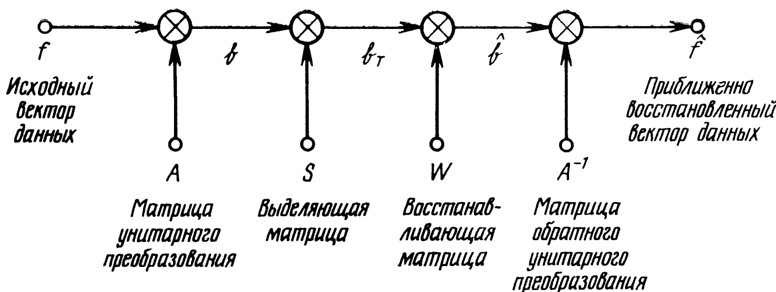


Рис. 23.3.1. Последовательность операций при экстраполяции спектра.

$$E\{(\mathbf{f} - \hat{\mathbf{f}})\mathscr{f}_T^{*T}\} = 0, \tag{23.3.1}$$

находим, что

$$E\{(\mathbf{f} - \mathbf{A}^{-1}\mathbf{W}\mathbf{S}\mathbf{f})(\mathbf{S}\mathbf{A}\mathbf{f})^{*T}\} = 0. \tag{23.3.2}$$

Отсюда сразу же следует оптимальное решение

$$\mathbf{W} = \mathbf{A}\mathbf{K}_f\mathbf{A}^{-1}\mathbf{S}^T[\mathbf{S}\mathbf{A}\mathbf{K}_f\mathbf{A}^{-1}\mathbf{S}^T]^{-1} \tag{23.3.3}$$

при условии, что фигурирующая здесь обратная матрица существует. Как показывает изучение строения матрицы $\mathbf{W}$, ее роль при формировании вектора $\hat{\mathscr{f}}$ сводится к копированию всех отобранных составляющих вектора $\mathscr{f}_T$, в то время как другие составляющие вектора $\hat{\mathscr{f}}$ восстанавливаются в виде линейных комбинаций составляющих вектора $\mathscr{f}_T$. По существу восстанавливающая матрица на основании значений известных коэффициентов спектра производит линейную экстраполяцию для определения неизвестных коэффициентов. Общепринятый способ действий, состоящий просто в замене отсеченных коэффициентов нулями перед выполнением обратного преобразования, можно рассматривать в соответствии с формулой (23.2.5) как другую форму экстраполяции, получаемую умножением $\mathscr{f}_T$ на $\mathbf{S}^T$.

Среднеквадратическая ошибка, получаемая в результате применения оптимальной восстанавливающей матрицы (23.3.3), вычисляется как

$$\mathscr{E}_O = \mathrm{tr}\,[\mathbf{K}_f(\mathbf{I}_N - \mathbf{W}\mathbf{S})]. \tag{23.3.4}$$

Графики на рис. 23.3.2 показывают среднеквадратическую ошибку оценки составляющих спектра в случаях преобразований Хаара и Адамара. В этих примерах предполагается, что кодированию

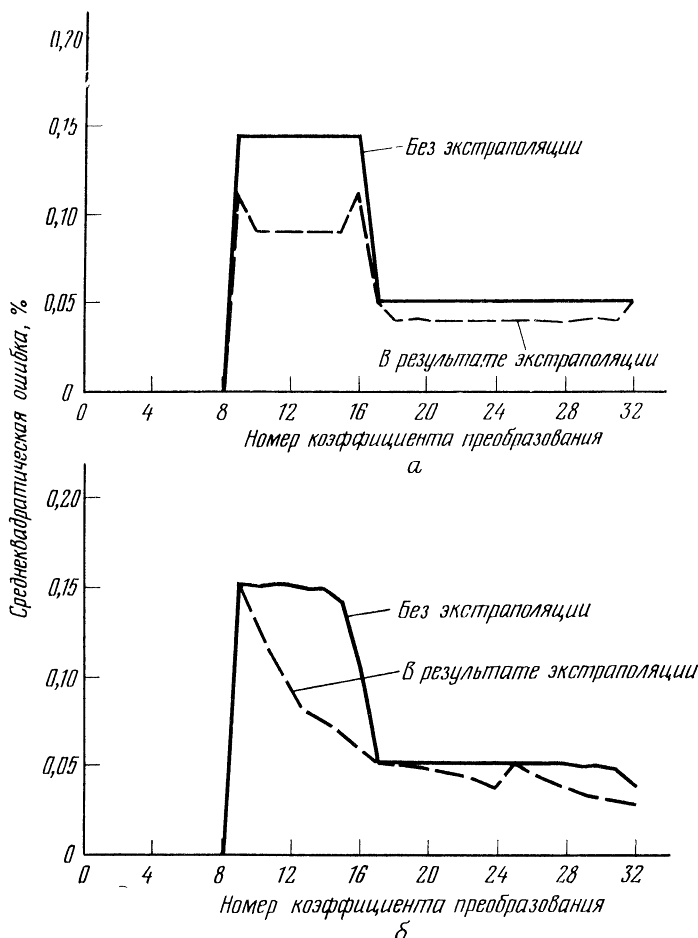


Рис. 23.3.2. Снижение среднеквадратической ошибки посредством экстраполяции спектра в случаях преобразований Хаара (*а*) и Адамара (*б*), примененных к марковскому процессу с коэффициентом корреляции $\rho = 0{,}95$. Кодирование с сохранением первых 8 из общего числа 32 коэффициентов преобразования.

подвергается марковский процесс первого порядка с коэффициентом корреляции $\rho = 0{,}95$. Эффективность экстраполяции спектра в процессе кодирования изображений изучалась также с помощью цифрового моделирования. Результаты этих экспериментов при-



Рис. 23.3.3. Примеры кодирования на основе преобразования Адамара с экстраполяцией спектра.

а — оригинал; *б* — зональный отбор с сохранением $^1/_4$ общего количества отсчетов, без экстраполяции; *в* — с применением экстраполяции; *г* — зональный отбор с сохранением $^1/_{10}$ общего количества отсчетов, без экстраполяции; *д* — с применением экстраполяции.

ведены на рис. 23.3.3. Исходное изображение, показанное на рис. 23.3.3, *а*, разделялось на блоки размером 16×16 элементов, подвергавшиеся преобразованию Адамара. Замена всех коэффициентов преобразования, находящихся за пределами низкочастотной зоны размером 5×5, нулями перед обратным преобразованием приводила к изображению, показанному на рис. 23.3.3, *б*. Результат восполнения коэффициентов за пределами этой зоны посредством экстраполяции в предположении, что изображение представляет собой реализацию двумерного марковского процесса первого порядка, показан на рис. 23.3.3, *в*. Фотоснимки демонстрируют повышение субъективного качества изображений в результате экстраполяции спектра.

Интерполяция спектра

Используя переменную шкалу квантования для различных коэффициентов преобразования, образующих вектор $\mathscr{f}$, каждый коэффициент кодируют отдельно. Формируемый в результате вектор **d** кодовых комбинаций определяет ячейку квантования. Такой же поэлементный характер носит обычно и процесс восстановления коэффициентов — k-я кодовая комбинация $d(k)$ определяет k-ю компоненту вектора $\hat{\mathscr{f}}(k), k = 1, 2, \ldots, N^2$. В случае интерполяции спектра к восстановлению вектора $\hat{\mathscr{f}}$ по вектору **d** подходят с точки зрения векторного (многомерного) квантования, описанного в разд. 6.2. Сначала, исходя из набора одномерных интервалов, указанных вектором **d**, определяют ячейку квантования D_j в N^2-мерном пространстве коэффициентов преобразования. Затем определяют восстанавливаемый вектор коэффициентов $\hat{\mathscr{f}}$ путем решения уравнения

$$\hat{\mathscr{f}} = \int_{D_j} \mathscr{f} p(\mathscr{f})\, d\mathscr{f} \Big/ \int_{D_j} p(\mathscr{f})\, d\mathscr{f}, \qquad (23.3.5)$$

где $p(\mathscr{f})$ — совместная плотность распределения вероятностей коэффициентов преобразования. За исключением ряда специальных случаев, решение уравнения (23.3.5) связано с большими трудностями. Приближенное решение в замкнутой форме может быть получено в случае гауссовой совместной плотности распределения при условии, что квантование производится не слишком грубо [40]. Рекуррентная процедура приближенного решения для того же гауссового случая, но с произвольными шкалами квантования была разработана Хунсом [39].

23.4. ГИБРИДНОЕ КОДИРОВАНИЕ С ИСПОЛЬЗОВАНИЕМ ПРЕОБРАЗОВАНИЯ И ДИКМ

Кодирование на основе преобразования и кодирование с предсказанием основаны на одном общем принципе — естественные корреляционные связи между элементами изображения разрушаются и параметры, полученные в результате декорреляции, по отдельности подвергаются квантованию и кодированию. При кодировании на основе преобразования этот принцип осуществляется путем отображения функции, описывающей исходное изображение, в новое векторное пространство, тогда как предсказание

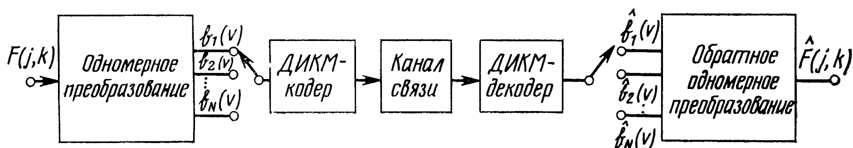


Рис. 23.4.1. Система гибридного кодирования с использованием преобразований и ДИКМ для передачи одноцветных изображений.

предусматривает обработку этой функции с помощью дифференцирования в контуре обратной связи. Каждый из методов имеет определенные преимущества и недостатки. В системах с наиболее экономным расходом двоичных цифр кодирование на основе двумерного преобразования обеспечивает более низкий уровень среднеквадратической ошибки воспроизведения, однако соответствующий кодер, как правило, значительно сложнее в практическом осуществлении по сравнению с кодером системы ДИКМ. Предложенный Хабиби [41] гибридный метод кодирования с использованием преобразований и ДИКМ сохраняет основные теоретические преимущества двумерных преобразований, но приводит к более простым в практическом осуществлении системам.

На рис. 23.4.1 приведена функциональная схема гибридного внутрикадрового кодирования. Вдоль каждой из строк в пределах блока $F(j, k)$, составленного из $N \times N$ элементов, производится одномерное преобразование с ядром $A(k, v)$, дающее последовательность коэффициентов

$$\mathscr{f}_j(v) = \sum_{k=0}^{N-1} F(j, k) A(k, v). \tag{23.4.1}$$

Коэффициенты преобразования поступают затем в буфер для последующей обработки с помощью ДИКМ. Эта обработка ведется в направлении столбцов исходного изображения и сводится к кодированию с предсказанием по предшествующему элементу. На прак-

тике можно применить поочередную обработку столбцов единственным кодером. В кодере формируется разностный сигнал

$$d(v) = f_j(v) - \hat{f}_j(v). \tag{23.4.2}$$

Оценка коэффициента преобразования $\hat{f}_j(v)$ формируется из оценки соответствующего коэффициента предыдущей строки, взятой с весом a_j:

$$\hat{f}_j(v) = a_j \hat{f}_{j-1}(v). \tag{23.4.3}$$

Квантованные разностные сигналы, полученные в результате ДИКМ, кодируются и перераспределяются во времени для передачи. На приемной стороне сначала декодер ДИКМ восстанавливает оценки $\hat{f}_j(v)$ коэффициентов преобразования. Затем в резуль-

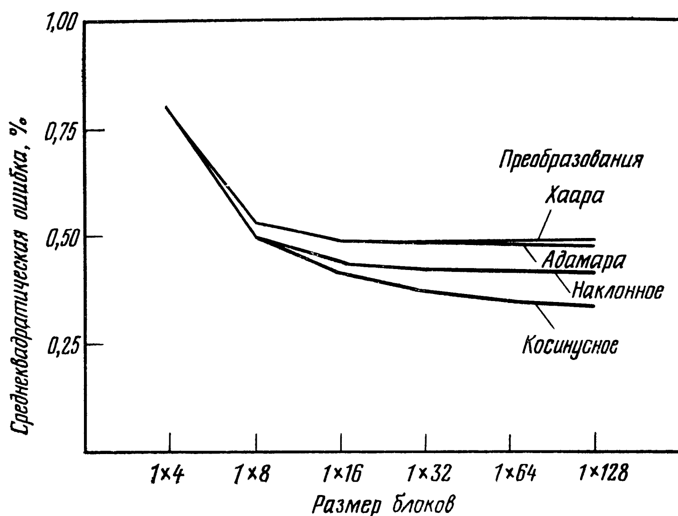


Рис. 23.4.2. Среднеквадратическая ошибка гибридного внутрикадрового кодирования с использованием преобразований и ДИКМ как функция размера блоков.

тате обратного преобразования с соответствующим ядром $B(k, v)$ получают восстановленное изображение

$$\hat{F}(j, k) = \sum_{v=0}^{N-1} \hat{f}_j(v) B(k, v). \tag{23.4.4}$$

Проектирование гибридного кодера для заданного преобразования будет заключаться в определении числа и расположения уровней квантования. Один из возможных подходов к проектированию состоит в том, чтобы в качестве статистической модели изобра-

жения взять двумерный случайный процесс с известной функцией ковариации. Статистические свойства коэффициентов одномерного преобразования моделируются при этом с помощью нормального



Рис. 23.4.3. Примеры гибридного внутрикадрового кодирования с использованием преобразований и ДИКМ при затрате 1 дв. ед. на элемент для блоков размером 1×16 элементов.

а — преобразование Адамара и ДИКМ; *б* — косинусное преобразование и ДИКМ; *в* — преобразование Карунена—Лоэва и ДИКМ.

распределения, а разностных сигналов, создаваемых ДИКМ, — с помощью экспоненциальных распределений, значения дисперсии которых непосредственно связаны со значениями дисперсии соответствующих коэффициентов преобразования. Затем для выбора пороговых уровней и уровней квантования, минимизирующих среднеквадратическую ошибку восстановления изображения, может быть применен алгоритм Макса.

На рис. 23.4.2 для некоторых преобразований графически показана зависимость среднеквадратической ошибки гибридного кодирования от размеров блока изображения. Как и следовало ожидать, наилучшие результаты дает косинусное преобразование, однако другие преобразования в случае гибридной системы лишь немногим ему уступают. На рис. 23.4.3 приведено несколько изо-

бражений, полученных методом гибридного кодирования. По субъективным оценкам, достаточно высокое качество изображений сохраняется при снижении цифрового потока вплоть до 1,0 дв.



Рис. 23.4.4. Примеры межкадрового адаптивного гибридного кодирования с использованием косинусного преобразования и ДИКМ. Последний 16-й по счету кадр блока. НСКП — нормированная среднеквадратическая ошибка.
а) 1,0 дв. ед./эл., НСКО = 0,022 %; *б*) 0,5 дв.ед./эл., НСКО= 0,067 %; *в*) 0,25 дв. ед./эл., НСКО = 0,186 %; *г*) 0,1 дв ед /эл., НСКО = 0,625 %.

ед./эл., а в некоторых практических случаях — даже до 0,5 дв. ед./эл.

Роуз и др. [37, 42] применили принцип гибридного кодирования на основе преобразования и ДИКМ для межкадрового кодирования. Один из вариантов состоял в выполнении одномерного преобразования вдоль каждой из строк и применении ДИКМ

к соответствующим коэффициентам в последовательности кадров. Лучших результатов удалось достичь при двумерном преобразовании в пределах кадра и использовании ДИКМ для коэффициентов в последовательности кадров. На рис. 23.4.4 приведены примеры адаптивного гибридного кодирования с применением косинусного преобразования к блокам размером 16×16 элементов.

23.5. КОДИРОВАНИЕ ПРИЗНАКОВ

Процесс зрительного восприятия той или иной сцены меньше всего напоминает количественный анализ яркости или координат цвета отдельных элементов изображения. Наблюдатель скорее стремится отыскать в изображении наиболее важные отличительные характеристики такого типа, как контуры или текстурные области, и образовать из них комбинации, поддающиеся распознаванию. Такие комбинации зрительных символов на дальнейших этапах обработки сравниваются с громадным фондом запасенных в памяти образцов с целью интерпретации сцены. При таком понимании процесса зрительного восприятия логично рассмотреть возможность эффективного кодирования характерных признаков изображения, а не его отдельных элементов. Были предложены методы кодирования линий уровня (изофот), контуров или областей с тем или иным типом текстуры.

23.5.1. КОДИРОВАНИЕ ЛИНИЙ УРОВНЯ

Квантованное изображение может рассматриваться как совокупность линий («ложных контуров»), разграничивающих области

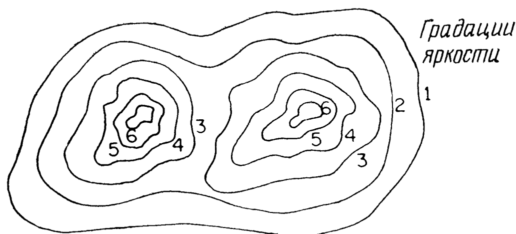
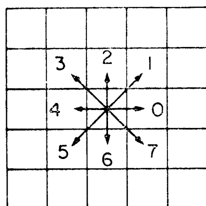


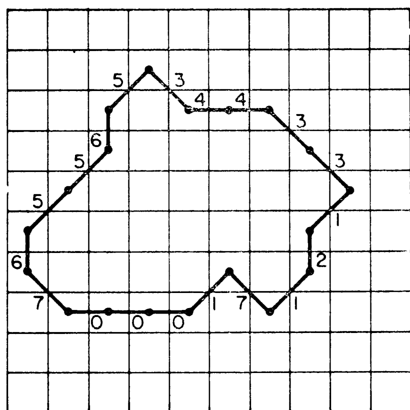
Рис. 23.5.1. Пример совокупности ложных контуров, разграничивающих области с постоянным уровнем яркости.

с одинаковым уровнем яркости. Каждому пороговому уровню в шкале квантования соответствует некоторый набор таких линий, определяющих в общей сложности изображение. Конфигурация линий уровня для типичного фрагмента изображения показана на рис. 23.5.1.

Уинц, Уилкинс и Гэттис [43—45] исследовали методы прослеживания линий уровня с помощью ЭВМ и эффективного кодирования этих линий для передачи по каналу связи. Кодирование линий уровня с целью сокращения полосы частот основывается на том, что такие линии относительно немногочисленны по сравнению с совокупностью элементов изображения. Кроме того, представляется более выгодным кодировать лишь очертания совокупностей элементов с одинаковым уровнем яркости, чем уровни элементов. Как показало моделирование с помощью ЭВМ, статистическое кодирование линий уровня дает возможность сократить полосу частот в отношении 7 : 1 в случае двухградационных изображений и в отношении 1,5 : 1 в случае изображений с числом градаций от 16 до 32 [46]. Однако кодирование линий уровня представляет собой достаточно сложный процесс и подходит для систем с невысокой скоростью передачи изображений.



а



б

Цепной код.
00017 12133 44356 5567

Рис. 23.5.2. Пример базовой операции цепного кодирования [48].

Остроумный метод кодирования линий, разграничивающих черные и белые области двухградационного изображения, разработан Фримэном [47, 48]. Этот метод не преследует цели наиболее эффективной минимизации количества двоичных цифр, затрачиваемых на описание линии. Однако получаемый в результате код (называемый цепным кодом) удобен для дальнейшей обработки изображения. Зная цепной код линии, можно достаточно просто вычислить различные параметры ограниченной ею области, такие, как площадь, периметр, максимальная высота и ширина и т. д. Пример цепного кода дан на рис. 23.5.2. Звенья цепи составляются прямолинейными отрезками, соединяющими соседние элементы изображения по вертикальному, горизонтальному и диагональному направлениям. Вертикально ориентированный отрезок, соединяющий 4 соседних элемента, будет иметь, например,

код 222. В примере, показанном на рис. 23.5.2, за начало отсчета в процессе цепного кодирования замкнутой линии выбран элемент, ближайший к левому нижнему углу кадра.

В случае набора вложенных друг в друга замкнутых границ может быть применена несложная модификация основной процедуры цепного кодирования. Сначала составляется обычный цепной код внешней границы. По окончании полного ее обхода в цепной код вводится кодовая группа 0401, указывающая на то, что последующая часть кода описывает «невидимую» линию, соединяющую первую границу со второй внутренней границей. По достижении внутренней границы в код вводится группа 0402, свидетельствующая об окончании невидимой линии, а затем кодирование продолжается обычным способом и т. д. Приставка 04 в кодовых группах, указывающих начало и окончание «невидимых» соединительных линий, соответствует той редко встречающейся ситуации, когда за смещением вправо сразу же следует смещение влево. Если такая ситуация все же возникает на самом деле, то она обозначается группой 0404. Приставка 04 используется также в сочетаниях с другими окончаниями в служебных кодовых группах, вводимых в различных специальных случаях, таких, как кодирование многоградационных или цветных изображений.

23.5.2. КОДИРОВАНИЕ КОНТУРОВ

Хорошо известно, что процесс интерпретации изображения наблюдателем в значительной степени основан на распознавании контуров. Далее, как показали исследования, зрительная система человека особо чувствительна к искажениям в расположении контуров, будучи, однако, менее критичной по отношению к искажениям в величине перепада яркости на контуре. Эти наблюдения послужили основой для разработки одномерных и двумерных схем кодирования контуров [49, 52].

В 1958 г. Шрайбер, Кнэпп и Кэй [49, 50] предложили метод кодирования контуров, известный под названием «синтетические высокие». При использовании этого метода видеосигнал, получаемый в результате развертки изображения, разделяется электрическими фильтрами на высокочастотную и низкочастотную компоненты. Такие компоненты для отрезка видеосигнала, содержащего импульс, показаны на рис. 23.5.3. Кодирующая система производит дискретизацию низкочастотной компоненты с частотой, значительно более низкой, чем частота Найквиста для исходного видеосигнала. К высокочастотной компоненте операция дискретизации непосредственно не применяется; производится выделение сигналов, соответствующих перепадам уровня на контурах, для передачи с помощью какого-либо подходящего кода. На приемной стороне при помощи линии задержки с отводами производится

синтез высокочастотной компоненты, которая добавляется затем к декодированной низкочастотной компоненте для восстановления исходного видеосигнала. Поскольку низкочастотная компонента не требует высокой частоты дискретизации, а перепады уровня относительно редки и допускают грубое квантование, становится возможным сократить полосу частот при передаче по цифровому каналу связи. На практике эффективность систем типа «синтетические высокие» ограничивается трудностями надежного выделения контуров из видеосигнала в присутствии шума.

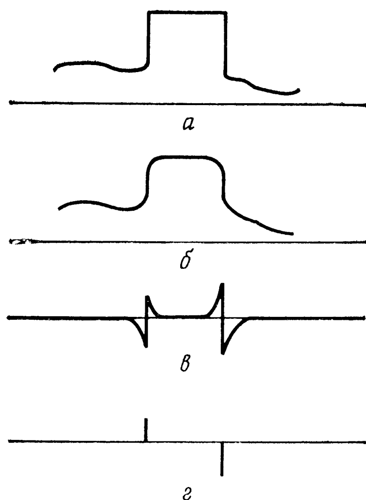


Рис. 23.5.3. Разделение видеосигнала на компоненты при кодировании методом «синтетические высокие».

а — исходный видеосигнал; *б* — низкочастотная компонента; *в* — высокочастотная компонента; *г* — импульсная компонента, передающая величину перепадов уровня на контурах.

Грэхэм [51] развил далее основную концепцию систем типа «синтетические высокие», придав ей двумерный характер. Низкочастотная компонента может быть получена с помощью двумерных фильтров нижних частот или путем грубой дискретизации с применением увеличенной апертуры. Для выделения контурных точек предусматривается процедура вычисления градиента или лапласиана с последующим объединением таких точек в связные границы. Как показало моделирование, средняя затрата двоичных цифр на передачу изображений с небольшим количеством деталей может быть снижена в такой системе примерно до 0,25 дв. ед./эл. Основной недостаток системы заключается в сложности ее практического осуществления.

23.5.3. КОДИРОВАНИЕ ТЕКСТУР

Многие изображения могут быть расчленены на отдельные области, каждая из которых имеет в достаточной мере одинаковую текстуру. Кроме того, возможно воспроизведение «синтетической» текстуры в границах области. Таким образом, представляют интерес измерения некоторых параметров текстуры в каждой области расчлененного изображения и последующий синтез текстуры в этих областях. Пример такой обработки изображения с помощью ЭВМ показан на рис. 23.5.4. В настоящее время ведутся исследования по вопросам эффективного анализа и синтеза текстур.

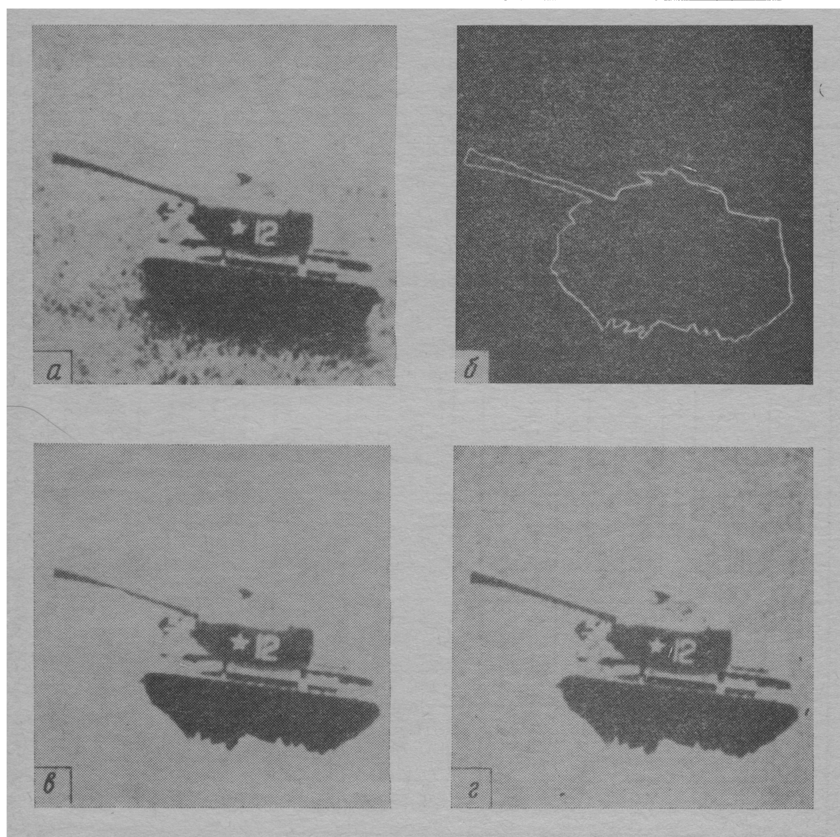


Рис. 23.5.4. Пример кодирования текстуры.

а — исходное изображение «танк»; *б* — очертание объекта; *в* — восстановленный объект на бестекстурном сером фоне; *г* — восстановленный объект на текстурированном фоне.

23.6. КОДИРОВАНИЕ С ПОМОЩЬЮ НАБОРА СИМВОЛОВ

В результате кодирования на основе преобразования изображение представляется в виде взвешенной комбинации двумерных функций, составляющих некоторый унитарный базис. Этот подход можно расширить, рассматривая в качестве средства для представления изображения также взвешенные комбинации двумерных образов или символов из наборов более общего вида.

23.6.1. КОДИРОВАНИЕ БЛОКОВ ПО СЛОВАРЮ ОБРАЗОВ

Одну из форм кодирования двухградационных изображений с помощью набора символов изучали Альтемус и Шэпхорст [53]. Изображение разбивается на блоки размером 5×5 элементов,

и каждый блок сравнивается с двухградационными образцами того же формата, составляющими кодовый словарь. Этот словарь из 31 образца показан на рис. 23.6.1. Если для данного блока в словаре имеется эквивалент, то передается соответствующая пятиразрядная кодовая комбинация. Если же эквивалент не обнаружен, то посылается специальная пятиразрядная кодовая комбинация, после чего производится полная поэлементная передача блока. При передаче слабо заполненных факсимильных изображений получено сокращение полосы частот в отношении 10 : 1 [53]. Распространение этого метода на случай многоградационных изображений вряд ли целесообразно из-за неизбежного расширения словаря образов.

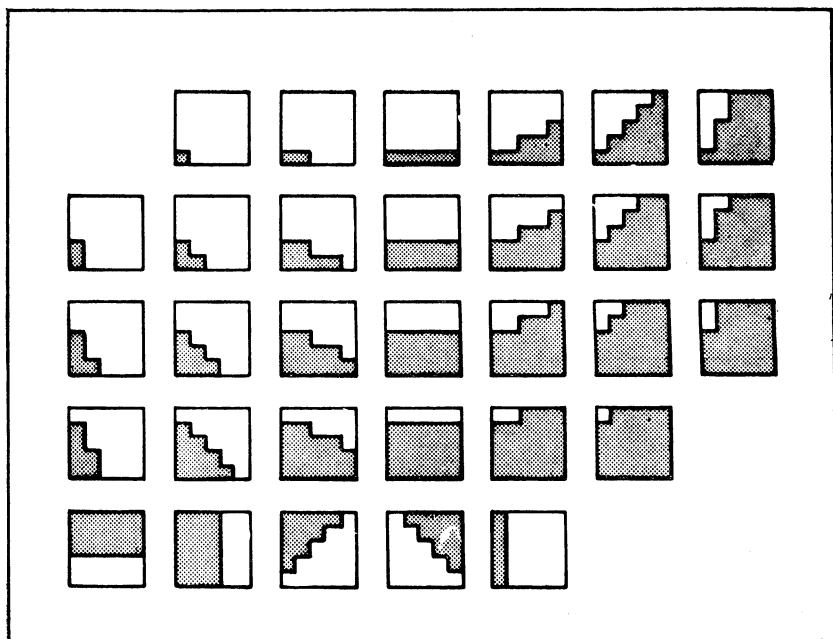


Рис. 23.6.1. Словарь двухградационных образов.

23.6.2. КОДИРОВАНИЕ НА ОСНОВЕ СИНГУЛЯРНОГО РАЗЛОЖЕНИЯ

Сингулярное разложение матриц (см. разд. 5.2) может служить средством для кодирования изображений. Сингулярное разложение матрицы **F**, описывающей исходное

изображение, определяется соотношениями (5.2.1) и (5.2.3):

$$\mathbf{\Lambda}^{1/2} = \mathbf{U}^T\mathbf{F}\mathbf{V}, \quad (23.6.1а)$$

$$\mathbf{F} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}^T, \quad (23.6.1б)$$

где $\mathbf{U}$ и $\mathbf{V}$ — унитарные матрицы, составленные соответственно из собственных векторов матриц $\mathbf{F}\mathbf{F}^T$ и $\mathbf{F}^T\mathbf{F}$, а $\mathbf{\Lambda}$ — диагональная матрица, элементами которой являются соответствующие собственные значения. Преобразование (23.6.1б) может рассматриваться как представление блока F с помощью базисного набора двумерных образов $\mathbf{u}_j\mathbf{v}_j^T$, получаемых из вектор-столбцов $\mathbf{u}_j$ и $\mathbf{v}_j$ матриц $\mathbf{U}$ и $\mathbf{V}$. Согласно формуле (5.2.6)

$$\mathbf{F} = \sum_{j=1}^{N} \lambda^{1/2}(j)\, \mathbf{u}_j\mathbf{v}_j^T, \quad (23.6.2)$$

где $\lambda(j)$ — диагональные элементы матрицы $\mathbf{\Lambda}$. Для большинства изображений естественного происхождения последовательность сингулярных значений $\lambda^{1/2}(j)$ убывает очень быстро, что позволяет ограничиться в разложении (23.6.2) малым числом членов без ущерба для точности восстановления. Эта операция аналогична зональному отбору коэффициентов при кодировании посредством унитарных преобразований. Отобранные сингулярные значения должны быть подвергнуты квантованию и кодированию для передачи по каналу связи. Кодированию подлежат также собственные векторы $\mathbf{u}_j$ и $\mathbf{v}_j$, поскольку они определяются изображением.

Метод кодирования на основе сингулярного разложения был проверен Эндрюсом и Паттерсоном [54] в условиях, когда изображение разбивалось на блоки размером 16×16 элементов, а кодирование сингулярных значений и собственных векторов производилось в расчете на минимизацию среднеквадратической ошибки. Как показало моделирование, при передаче с погрешностью около 1 % цифровой поток может составлять 2 дв. ед./эл. По сравнению с этим результатом преобразование обычного типа имеет определенное преимущество.

ЛИТЕРАТУРА

1. Gardenhire L. W., Redundancy Reduction, the Key to Adaptive Telemetry, National Telemetering Conference, Los Angeles, California, 1964.
2. Kortman C. M., Redundancy Reduction: A Practical Method of Data Compression, *Proc. IEEE*, **55**, 3, 253—263 (March 1967). [Имеется перевод: Котман. Сокращение избыточности как практический метод сжатия данных. — ТИИЭР, тематический выпуск. «Сокращение избыточности», 1967, т. 55, № 3, с. 8—21.]
3. Hochman D., Katzman H., Weber D. R., Application of Redundancy Reduction to Television Bandwidth Compression, *Proc. IEEE*, **55**, 3, 263—267

(March 1967). [Имеется перевод: Хочман, Кацман, Вебер. Сжатие полосы телевизионных сигналов за счет сокращения избыточности. — ТИИЭР, тематический вып. «Сокращение избыточности», 1967, т. 55, № 3, с. 21—35.]

4. Andrews C. A., Davies J. M., Schwarz G. R., Adaptive Data Compression, *Proc. IEEE*, **55**, 3, 267—277 (March 1967). [Имеется перевод: Эндрюс, Дэвис, Шварц. Адаптивное сжатие данных. — ТИИЭР, тематический вып. «Сокращение избыточности», 1967, т. 55, № 3, с. 25—38.]
5. Kortman C. M., Data Compression by Redundancy Reduction, *IEEE Spectrum*, **4**, 133—139 (March 1967).
6. Ehrman L., Analysis of Some Redundancy Removal Bandwidth Compression Techniques, *Proc. IEEE*, **55**, 3, 278—287 (March 1967). [Имеется перевод: Эрман. Анализ некоторых способов сокращения полосы частот путем устранения избыточности. — ТИИЭР, тематический вып. «Сокращение избыточности», 1967, т. 55, № 3, с. 38—48.]
7. Davisson L. D., Data Compression Using Straight Line Interpolation, *IEEE Trans. Inf. Theory*, **IT-14**, 3, 390—394 (May 1968).
8. Sheldahl S. A., Comparison of Hardware Requirements for Polynomial Data Compressions, Proceedings National Telemetering Conference, Washington, D. C., 1968.
9. Gabor D., Hill P. C. J., Television Bandwidth Compression by Contour Interpolation, *Proc. IEE*, **108**, Part B, 39, 634, 303—315 (May 1961).
10. Andrews H. C., Digital Interpolation of Discrete Images, *IEEE Trans. Computers*, **C-25**, 2, 196—202 (February 1976).
11. Andrews H. C., Pratt W. K., Fourier Transform Coding of Images, Hawaii International Conference on System Science, January 1968, pp. 677—679.
12. Anderson G. B., Huang T. S., Piecewise Fourier Transformation for Picture Bandwidth Compression, *IEEE Trans. Commun.*, COM-**20**, 3, 488—491 (June 1972).
13. Andrews H. C., Kane J., Pratt W. K., Hadamard Transform Image Coding, *Proc. IEEE*, **57**, 1, 58—68 (January 1969). [Имеется перевод: Пратт, Кэйн, Эндрюс. Кодирование изображений посредством преобразования Адамара.— ТИИЭР, 1969, т. 57, № 1, с. 66—77.]
14. Pratt W. K., Andrews H. C., Application of Fourier-Hadamard Transformation to Bandwidth Compression, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordon and Breach, New York, 1972, pp. 515—554.
15. Woods J. W., Huang T. S., Picture Bandwidth Compression by Linear Transformation and Block Quantization, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordong and Breach, New York, 1972, pp. 555—573.
16. Habibi A., Wintz P. A., Image Coding by Linear Transformation and Block Quantization, *IEEE Trans. Commun. Tech.*, COM-**19**, 1, 50—63 (February 1971).
17. Andrews H. C., Computer Techniques in Image Processing, Academic Press, New York, 1970.
18. Rao K. R., Narasimhan M. A., Revuluri K., Image Data Processing by Hadamard-Haar Transforms, *IEE* Trans. *Computers*, **C-23**, 9, 888—896 (September 1975).
19. Enomoto H., Shibata K., Orthogonal Transform Coding System for Television Signals, *IEEE Trans. Electromagnetic Compatibility*, Special Issue on Walsh Functions, **EMC-13**, 3, 11—17 (August 1971).
20. Pratt W. K., Chen W. H., Welch L. R., Slant Transform Image Coding, *IEEE Trans. Commun.*, **COM-22**, 1075—1093 (August 1974).
21. Ahmed N., Natarajan T., Rao K. R., On Image Processing and a Discrete Cosine Transform, *IEEE Trans. Computers*, **C-23**, 1, 90—93 (January 1974).

22. Jain A. K., A Fast Karhunen-Loeve Transform for Finite Discrete Images, Proceedings National Electronics Conference, Chicago, Illinois, October 1974, pp. 322—328.
23. Andrews H. C., Pratt W. K., Transform Image Coding, in: Proceedings Computer Processing in Communications, Polytechnic Press, New York, 1969, pp. 63—84.
24. Wintz P. A., Transform Picture Coding, *Proc. IEEE*, **60**, 7, 809—823 (July 1972). [Имеется перевод: Уинтц. Кодирование изображений посредством преобразований. — ТИИЭР, тематический вып. «Обработка изображений при помощи цифровых вычислительных машин», 1972, т. 60, № 7, с. 69—84.]
25. Pratt W. K., Transform Coding Spectrum Extrapolation, Proceedings Hawaii Systems Science Conference, January 1974.
26. Kramer H. P., Mathews M. V., A Linear Coding for Transmitting a Set of Correlated Signals, *IRE Trans. Inf. Theory*, **IT-2**, 41—46 (September 1956).
27. Huang J. J. Y., Schulteiss P. M., Block Quantization of Correlated Gaussian Random Variables, *IEEE Trans. Commun. Syst.*, **CS-11**, 3, 289—286 (September 1963).
28. Reis J. J., Lynch R. T., Butman J., Haar Transform Video Bandwidth Reduction System for RPVs, Proceedings SPIE Conference on Advances in Image Transmission Techniques, San Diego, California, August 1976.
29. Andrews H. C., Tescher A. G., The Role of Adaptive Phase Coding in Two and Three Dimensional Fourier and Walsh Image Compression, Proceedings Walsh Function Symposium, Washington, D. C., March 1974.
30. Tasto M., Wintz P. A., A Bound on the Rate-Distortion Function and Application to Images, *IEEE Trans. Inf. Theory*, **IT-18**, 1, 150—159 (January 1972).
31. Tasto M., Wintz P. A., Image Coding by Adaptive Block Quantization, *IEEE Trans. Commun. Tech.*, **COM-19**, 6, 957—972 (December 1971).
32. Chen W., Adaptive Coding of Color Images Using Cosine Transform, Proceedings International Communications Conference, Philadelphia, June 1976.
33. Cox R. V., Tescher A. G., Channel Rate Equalization Techniques for Adaptive Transform Coders, Proceedings SPIE Conference on Advances in Image Transmission Techniques, San Diego, August 1976.
34. Pratt W. K., Spatial Transform Coding of Color Images, *IEEE Trans. Commun. Tech.*, **COM-19**, 6, 980—982 (December 1971).
35. Enomoto H., Shibata K., Orthogonal Transform Coding System for Television Signals, Applications of Walsh Function Symposium, Washington, D. C., 1971, pp. 11—17.
36. Rose J. A., Pratt W. K., Theoretical Performance Models for Interframe Transform and Hybrid Transform/DPCM Coders, Proceeding SPIE Conference on Advances in Image Transmission Techniques, San Diego, August 1976.
37. Rose J. A. et al., Interframe Transform Coding and Predictive Coding Methods, IEEE International Communications Conference, San Francisco, California, June 1975.
38. Pratt W. K., Transform Image Coding Spectrum Extrapolation, Proceedings Hawaii Systems Science Conference, January 1974.
39. Huhns M. N., Optimum Restoration of Quantized Correlated Signals, University of Southern California, Image Processing Institute, USCIPI 600, August 1975.
40. Curry R. E., Estimation and Control with Quantized Measurements, MIT Press, Cambridge, Mass., 1970.
41. Habibi A., Hybrid Coding of Pictorial Data, *IEEE Trans. Computers*, **COM-22**, 5, 614—624 (May 1974).
42. Rose J. A., Pratt W. K., Robinson G. S., Interframe Cosine Transform Image Coding, *IEEE Trans. Commun.*, **COM-25**, 11 (November 1977).
43. Gattis J. L., Wintz P. A., Automated Techniques for Data Analysis and Transmission, Purdue University, School of Engineering, Report TR-EE-37, August 1971.

44. Wilkins L. C., Wintz P. A., A Contour Tracing Algorithm for Data Compression of Two-Dimensional Data, Purdue University, School of Engineering, Report TR-EE-69-3, January 1969.
45. Wintz P. A., Wilkins L. C., Studies on Data Compression, Part I: Picture Coding by Contours, Part II: Error Analysis of Run-Length Codes, Purdue University, School of Engineering, Report TR-EE-70-17, September 1970.
46. Wilkins L. C., Wintz P. A., Image Coding by Coding Contours, Proceedings International Conference on Communications, San Francisco, Vol. 1, 1970.
47. Freeman H., On the Encoding of Arbitrary Geometric Configurations, *IRE Trans. Electronic Computers*, **EC-10**, 2, 260—268 (June 1961).
48. Freeman H., Boundary Encoding and Processing, in: Picture Processing, and Psychopictorics, Lipkin B. S., Rosenfeld A., Eds., Academic Press, New York, 1970, pp. 241—266.
49. Schreiber W. F., Knapp C. F., TV Bandwidth Reduction by Digital Coding, *IRE National Convention Record*, **6**, Part 4, 88—89 (1958).
50. Schreiber W. F., Knapp C. F., Kay N. D., Synthetic Highs: An Experimental TV Bandwidth Reduction System, *J. Soc. Motion Picture and Television Engineers*, **68**, 525—537 (August 1959).
51. Graham D. N., Image Transmission by Two-Dimensional Contour Coding, *Proc. IEEE*, **55**, 3, 336—346 (March 1967). [Имеется перевод: Грэхэм. Передача изображений посредством кодирования двумерных контуров. — ТИИЭР, тематический вып. «Сокращение избыточности», 1967, т. 55, № 3, с. 102—114.]
52. Schreiber W. F., Huang T. S., Tretiak O. J., Contour Coding of Images, in: Picture Bandwidth Compression, Huang T. S., Tretiak O. J., Eds., Gordon and Breach, New York, 1972, pp. 443—448.
53. Altemus W. C., Schaphorst R. A., Data Conversion Equipment, Philco Corporation Report No. ASD TDR 62-392, AD 275 514, May 1962.
54. Andrews H. C., Patterson C. L., Singular Value Decomposition (SVD) Image Coding, *IEEE Trans. Commun.*, **COM-24**, 4, 425—432 (April 1976).

Глава 24

АНАЛИЗ ЭФФЕКТИВНОСТИ КОДИРОВАНИЯ ИЗОБРАЖЕНИЙ

Эта глава посвящена оценке важнейших методов кодирования изображений. Приведены результаты теоретического и экспериментального изучения действия помех канала связи на процесс передачи изображения. Для различных методов кодирования установлены теоретические границы расхода двоичных цифр на элемент изображения в зависимости от уровня искажений. Эти границы сопоставлены с оценками и экспериментальными данными, характеризующими эффективность ряда практических систем кодирования изображений.

24.1. СТАТИСТИЧЕСКИЕ МОДЕЛИ КАНАЛА СВЯЗИ КАК ИСТОЧНИКА ПОМЕХ [1—4]

Поскольку изображения в процессе передачи по каналу связи подвергаются воздействию помех, следует рассмотреть последствия такого воздействия и возможности борьбы с помехами путем надлежащего кодирования. Проблема помехоустойчивости имеет различные аспекты и требует учета таких обстоятельств, как тип и целевое назначение передаваемых изображений, критерий оценки качества (степени искажений) изображения, методы кодирования, выбранные для согласования с источником сообщений и с каналом связи, и, наконец, статистические свойства канала как источника помех.

Для оценки характеристик цифрового канала связи применяются различные статистические модели. Наиболее простая и широко используемая модель — это двоично-симметричный канал (ДСК), меняющий в результате действия помех значение переданного двоичного разряда на обратное с вероятностью p и сохраняющий это значение неизменным с вероятностью $1 - p$. Модель включает предположение о том, что ошибки в двоичных разрядах возникают независимо друг от друга. На практике приходится также сталкиваться с каналами, где ошибки возникают пачками (сериями). Ради упрощений при изучении таких каналов часто предполагается, что пачки ошибок имеют фиксированную длину и возникают независимо друг от друга. В рамках модели ДСК это означает, что пачка ошибок длиной M двоичных разря-

дов появляется по предположению с вероятностью p. Дальнейшее рассмотрение воздействия помех на процесс передачи изображения ограничивается простым случаем двоично-симметричного канала.

Пусть двоичное число k, состоящее из N разрядов, представляет результат равномерного квантования уровня некоторой компоненты изображения. На вход двоично-симметричного канала поступает, таким образом, число

$$k = \sum_{i=0}^{N-1} k_i 2^i, \quad k_i \in \{0, 1\}. \tag{24.1.1}$$

Число на выходе канала определяется как

$$h = \sum_{i=0}^{N-1} h_i 2^i, \quad h_i \in \{0, 1\}, \tag{24.1.2}$$

а ошибка, возникшая при передаче по каналу, — как $\mathscr{E} = h - k$.

При изучении действия помех основной интерес представляют среднее смещение яркости (или цвета) и дисперсия этого смещения, но не среднее значение яркости (цвета) элемента изображения. Важны, следовательно, такие статистические характеристики рассматриваемой компоненты изображения, как первый и второй условные моменты распределения для значений двоичного числа на выходе канала.

Условное распределение вероятностей значений выходного числа h относительно входного числа k представляет собой геометрическое распределение

$$P\{h \mid k\} = p^{N-M\{h,\,k\}}(1-p)^{M\{h,\,k\}}, \tag{24.1.3}$$

где $M\{h, k\}$ есть количество совпадающих разрядов чисел h и k, определяемое как

$$M\{h, k\} = \sum_{i=0}^{N-1} [h_i k_i + (1-h_i)(1-k_i)]. \tag{24.1.4}$$

В отсутствие ошибок при передаче по каналу $P\{h \mid k\} = \delta(h - k)$; если же $p = 0{,}5$, то значения выходного числа распределены равномерно с плотностью $P\{h \mid k\} = 2^{-N}$. Последний случай представляет интерес при установлении верхнего предела для смещения яркости.

Условное среднее значение выходного числа может быть выражено через условные средние для каждого разряда этого числа следующим образом:

$$E\{\mathrm{h} \mid k\} = \sum_{i=0}^{N-1} 2^i E\{h_i \mid k_i\}. \tag{24.1.5}$$

В процессе передачи значение k_i двоичного разряда заменяется на $1 - k_i$ с вероятностью p и сохраняется без изменения с вероят-

ностью $1 - p$. Таким образом, условное среднее для данного разряда на выходе канала есть

$$E\{h_i \mid k_i\} = p(1 - k_i) + (1 - p)k_i. \qquad (24.1.6)$$

Суммирование по всем i дает

$$E\{h \mid k\} = (1 - 2p)k + p(2^N - 1). \qquad (24.1.7)$$

Второй условный момент для значения выходного числа определяется как

$$E\{h^2 \mid k\} = E\left\{\left[\sum_{i=0}^{N-1} h_i 2^i\right]^2 \middle| k\right\} =$$

$$= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} 2^{i+j} E\{h_i h_j \mid k_i k_j\}. \qquad (24.1.8)$$

При $i \neq j$

$$E\{h_i h_j \mid k_i k_j\} = (1 - p)^2 k_i k_j + (1 - p) p k_i (1 - k_j) +$$

$$+ p(1 - p)(1 - k_i) k_j + p^2 (1 - k_i)(1 - k_j), \qquad (24.1.9)$$

а при $i = j$

$$E\{h_i^2 \mid k_i\} = (1 - p) k_i^2 + p(1 - k_i)^2. \qquad (24.1.10)$$

Суммирование по всем разрядам дает в результате

$$E\{h^2 \mid k\} = [k(1 - 2p) + p(2^N - 1)]^2 + \frac{4N - 1}{3} p(1 - p), \qquad (24.1.11)$$

и условная дисперсия оказывается равной

$$\sigma^2\{h \mid k\} = [(4^N - 1)/3]\, p(1 - p). \qquad (24.1.12)$$

Кроме простого двоичного кода существуют и другие коды, которыми можно было бы воспользоваться для представления исходного сигнала. Для большинства приложений наилучшим является код, минимизирующий дисперсию ошибки (24.1.12). Однако Хуанг и др. [3], изучавшие этот вопрос, не обнаружили среди различных кодов, включая код Грэя, такого, который уменьшал бы дисперсию ошибки по сравнению с простым двоичным кодом.

24.2. ДЕЙСТВИЕ ПОМЕХ ПРИ ПЕРЕДАЧЕ ОДНОЦВЕТНЫХ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ ИКМ

Для анализа действия помех при передаче сигнала яркости одноцветного изображения удобно воспользоваться схемой, приведенной на рис. 24.2.1. Нормированный сигнал яркости Y, возможные значения которого заключены в интервале [0, 1],



Рис. 24.2.1. Функциональная модель процесса передачи сигнала яркости, служащая для анализа действия помех канала связи.

в процессе равномерного квантования приобретает значения от 0 до $2^N - 1$ и округляется до целого, ближайшего к

$$Y_T = Y(2^N - 1). \tag{24.2.1}$$

Значение Y'_T, принятое в результате прохождения соответствующей кодовой комбинации по двоично-симметричному каналу, меняется затем согласно соотношению

$$Y' = Y'_T/(2^N - 1). \tag{24.2.2}$$

Условный момент n-го порядка величины Y' относительно переданной компоненты яркости Y_T может быть определен непосредственным расчетом по формуле

$$E\{(Y')^n \mid Y_T\} = \sum_{Y'_T=0}^{2N-1} [Y'(Y'_T)]^n P\{Y'_T \mid Y_T\}, \tag{24.2.3}$$

где условная плотность распределения $P\{Y'_T \mid Y_T\}$ дается выражением (24.1.3). Однако, поскольку прямая и обратная операции масштабирования линейны, условные моменты смещения яркости можно легко выразить через условные моменты, определенные в разд. 24.1 в качестве характеристик ДСК. Условное среднее смещение сдвига яркости есть

$$E\{Y' \mid Y_T\} = E\{Y'_T \mid Y_T\}/(2^N - 1). \tag{24.2.4}$$

Согласно формуле (24.1.7),

$$E\{Y' \mid Y_T\} = [(1-2p)Y_T/(2^N-1)] + p. \qquad (24.2.5)$$

Аналогично с помощью выражения (24.1.12) определяется дисперсия сдвига яркости как

$$\sigma^2\{Y' \mid Y_T\} = \frac{4^N-1}{3}\frac{p(1-p)}{(2^N-1)^2} = \frac{2N+1}{3}p(1-p). \qquad (24.2.6)$$

На рис. 24.2.2 показана графически зависимость среднего смещения яркости от значения кодированного уровня яркости Y_T и соответствующего входного уровня яркости Y. Чем ближе к середине воспроизводимого диапазона яркости, тем более от-

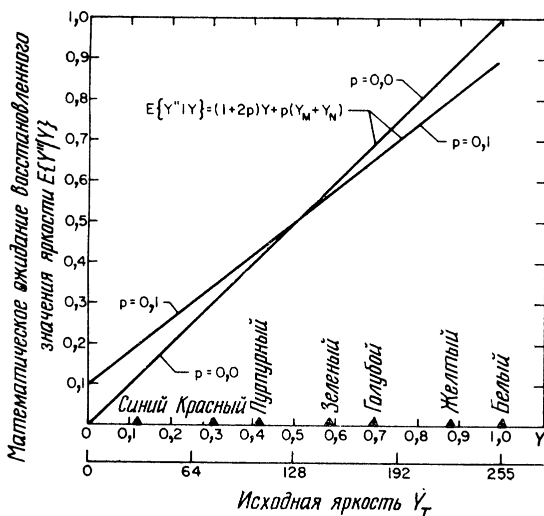


Рис. 24.2.2. Смещение яркости под воздействием помех при передаче сигнала яркости по двоично-симметричному каналу.

даляется среднее смещение яркости от тех значений, которые могли бы вызвать значительный контраст.

Рис. 24.2.3 иллюстрирует на нескольких примерах воздействие помех канала связи на процесс передачи одноцветного изображения. «Соль и перец» — так часто называют зрительный эффект, создаваемый помехами в системе ИКМ, поскольку они проявляются в виде черных точек на светлых участках изображения и белых точек — на темных участках. Субъективное качество изображения обычно остается приемлемым при повышении частоты ошибок вплоть до 10^{-3} [1, 5].

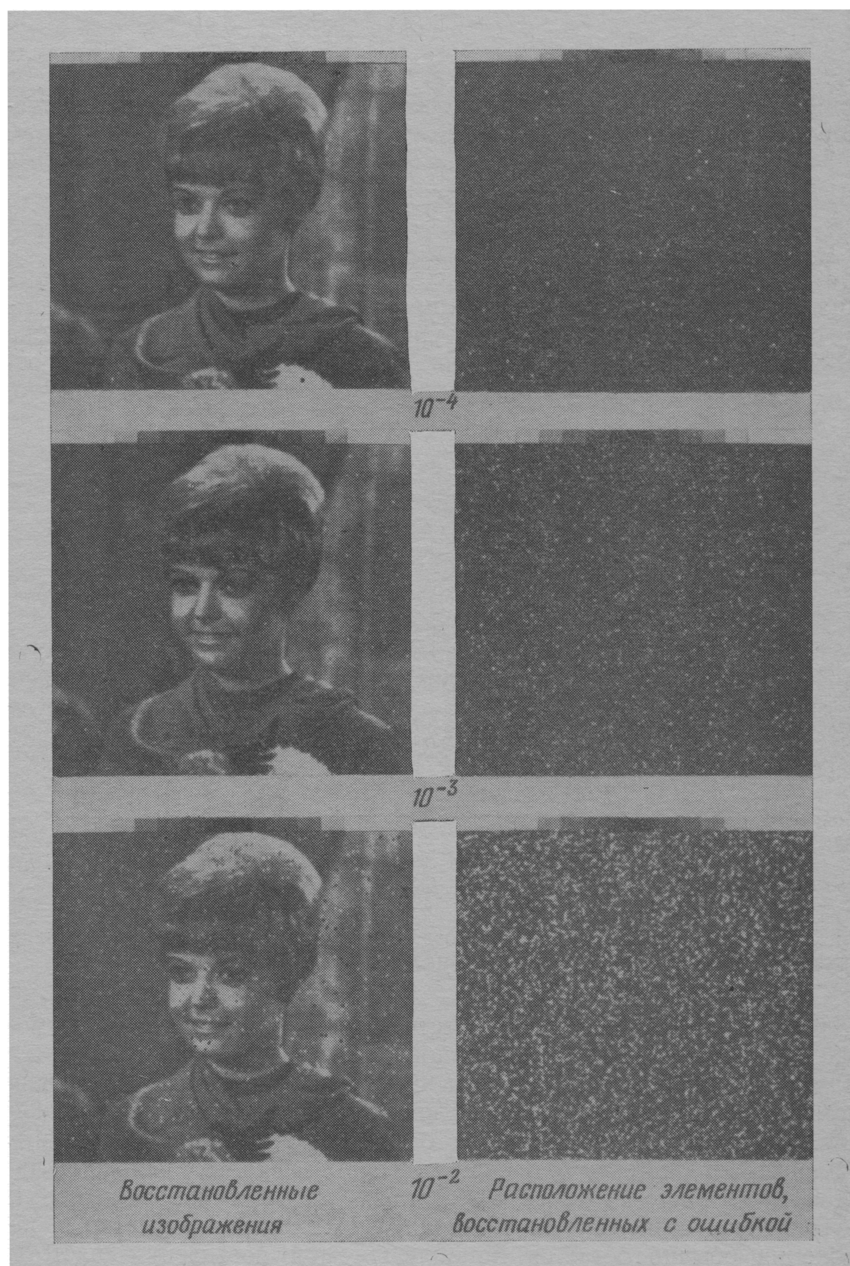


Рис. 24.2.3. Действие помех канала связи в системе ИКМ.

24.3. ДЕЙСТВИЕ ПОМЕХ ПРИ ПЕРЕДАЧЕ ЦВЕТНЫХ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ ИКМ [4]

Воздействие помех канала связи на процесс передачи цветных изображений рассматривается для двух случаев: передачи координат цвета и передачи яркости в сочетании с двумя координатами цветности. В обоих случаях действие помех оценивается по тем искажениям, которым подвергаются яркость Y и координаты u, v равноконтрастного цветового графика (РЦГ). Структура системы передачи для каждого из указанных случаев показана на рис. 24.3.1.

24.3.1. ПЕРЕДАЧА КООРДИНАТ ЦВЕТА

В системе, показанной на рис. 24.3.1, *а*, нормированные первичные координаты цвета R, G, B, значения каждой из которых заключены в интервале [0, 1], подвергаются линейному преобразованию для формирования передаваемых сигналов цвета

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \qquad (24.3.1)$$

причем коэффициенты m_{ij} фиксированы. Каждая из компонент T_i с диапазоном изменения от T_{iN} до T_{iM} подвергается затем в процессе равномерного квантования масштабированию и округлению до целого числа:

$$T_{iT} = [(T_i - T_{iN})/(T_{iM} - T_{iN})]\left(2^N - 1\right). \qquad (24.3.2)$$

Принятая компонента T'_{iT} сигнала цвета в результате обратной операции масштабирования принимает значение

$$T'_i = [(T_{iM} - T_{iN})/\left(2^N - 1\right)]T'_{iT} + T_{iN}. \qquad (24.3.3)$$

Далее производится преобразование координат цвета, дающее принятые координаты цвета R', G', B'. Помехи в канале связи могут привести к расширению диапазона изменения этих величин по сравнению с интервалом [0, 1], воспроизводимым дисплеем. Поэтому необходимо ввести ограничение с тем, чтобы диапазон изменения координат цвета R'', G'', B'', полученных на выходе ограничителей, соответствовал физически воспроизводимым цветам.

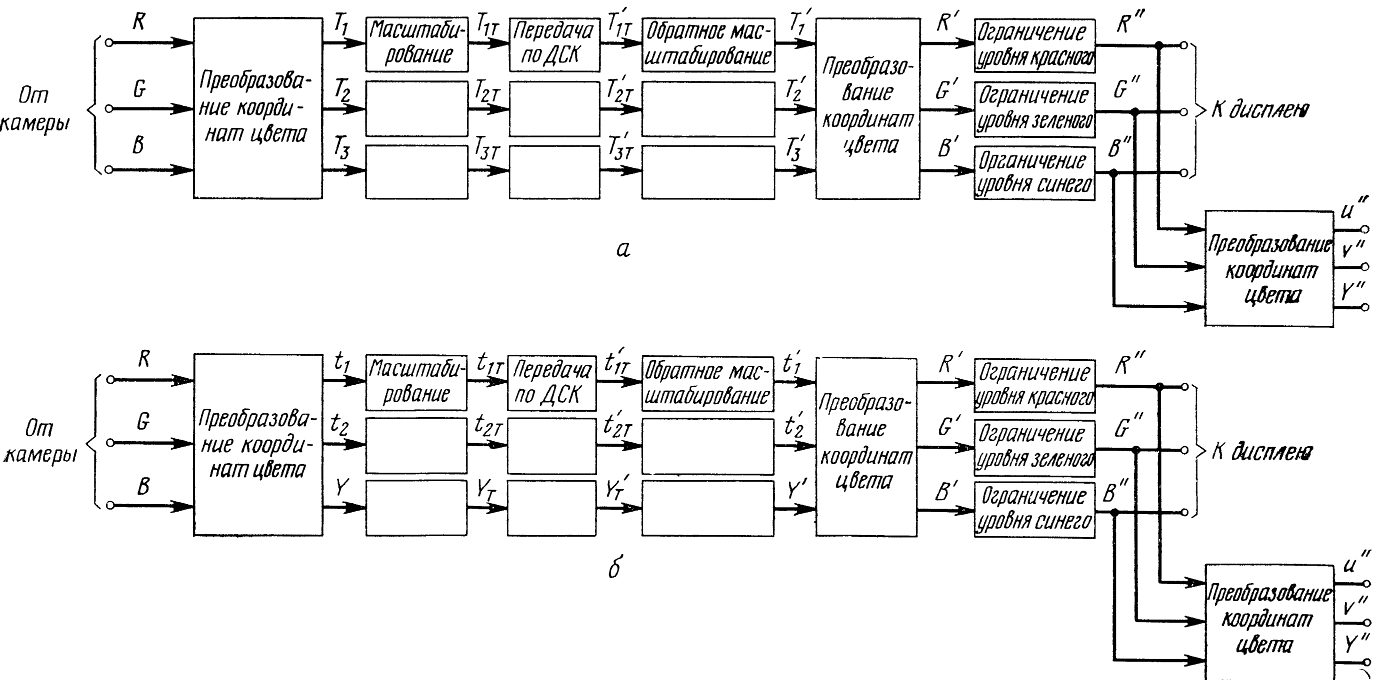


Рис. 24.3.1. Функциональные модели процесса передачи цветного изображения, используемые для анализа действия помех канала связи.

а — передача координат цвета T_1, T_2, T_3; *б* — передача координат цветности Y, t_1, t_2.

Схема на рис. 24.3.1 преобразует координаты цвета (R'', G'', B''), поступающие на дисплей, в координаты яркости Y'' и цветности (u'', v''):

$$Y'' = \lambda_1 R'' + \lambda_2 G'' + \lambda_3 B'', \qquad (24.3.4)$$

$$u'' = (\eta_1 R'' + \eta_2 G'' + \eta_3 B'')/(\eta_4 R'' + \eta_5 G'' + \eta_6 B''), \qquad (24.3.5)$$

$$v'' = (\eta_7 R'' + \eta_8 G'' + \eta_9 B'')/(\eta_4 R'' + \eta_5 G'' + \eta_6 B''), \qquad (24.3.6)$$

где λ_k и η_k — постоянные, определяющие это преобразование.

Смещение цветов под воздействием помех определяется затем путем вычисления условных моментов для Y'', u'' и v''. Условный момент n-го порядка величины u'' относительно переданных компонент $\{M\} = \{T_{1T}, T_{2T}, T_{3T}\}$ может быть непосредственно вычислен по формуле

$$E\{(u'')^n \mid M\} = \sum_{T'_{1T}=0}^{2N-1} \sum_{T'_{2T}=0}^{2N-1} \sum_{T'_{3T}=0}^{2N-1} [u''\{T'_{1T}, T'_{2T}, T'_{3T}\}]^n \times$$
$$\times P\{T'_{1T} \mid T_{1T}\} P\{T'_{2T} \mid T_{2T}\} P\{T'_{3T} \mid T_{3T}\}. \qquad (24.3.7)$$

Аналогичные выражения получаются для $E\{(v'')^n \mid M\}$ и $E\{(Y'')^n \mid M\}$.

Условные распределения $P\{T'_{iT} \mid T_{iT}\}$ могут быть рассчитаны в соответствии с формулой (24.1.3), а значение функции u'', стоящей под знаками сумм (и аналогично v'' или Y''), может быть определено для любой комбинации значений ее переменных T_{iT}. К сожалению, при больших N прямое вычисление условных моментов связано со значительными затратами времени.

В случае канала с малой вероятностью ошибок условные распределения вероятностей $P\{T'_{iT} \mid T_{iT}\}$ характеризуются небольшим разбросом относительно своих средних. В этом случае вероятность того, что принятые координаты цвета R', G', B' выйдут за установленные для них пределы, будет достаточно мала. Поэтому можно приближенно найти условные средние для u'', v'' и Y'' [5, стр. 212], пренебрегая операцией ограничения и выражая u'' (аналогично v'') и Y'' непосредственно через принятые значения переменных T'_{iT}. Таким образом,

$$u'' \approx (a_1 T'_{1T} + a_2 T'_{2T} + a_3 T'_{3T} + a_4)/(a_5 T'_{1T} + a_6 T'_{2T} + a_7 T'_{3T} + a_8), \qquad (24.3.8)$$

$$v'' \approx (a_9 T'_{1T} + a_{10} T'_{2T} + a_{11} T'_{3T} + a_{12})/(a_5 T'_{1T} + a_6 T'_{2T} + a_7 T'_{3T} + a_8), \qquad (24.3.9)$$

$$Y'' \approx a_{13} T'_{1T} + a_{14} T'_{2T} + a_{15} T'_{3T} + a_{16}, \qquad (24.3.10)$$

где

$$a_1 = \beta_1 (T_{1M} - T_{1N}),$$
$$a_2 = \beta_2 (T_{2M} - T_{2N}),$$
$$a_3 = \beta_3 (T_{3M} - T_{3N}),$$
$$a_4 = (2^N - 1)(\beta_1 T_{1N} + \beta_2 T_{2N} + \beta_3 T_{3N}),$$
$$a_5 = \beta_4 (T_{1M} - T_{1N}),$$
$$a_6 = \beta_5 (T_{2M} - T_{2N}),$$
$$a_7 = \beta_6 (T_{3M} - T_{3N}),$$
$$a_8 = (2^N - 1)(\beta_4 T_{1N} + \beta_5 T_{2N} + \beta_6 T_{3N}),$$
$$a_9 = \beta_7 (T_{1M} - T_{1N}),$$
$$a_{10} = \beta_8 (T_{2M} - T_{2N}),$$
$$a_{11} = \beta_9 (T_{3M} - T_{3N}),$$
$$a_{12} = (2^N - 1)(\beta_7 T_{1N} + \beta_8 T_{2N} + \beta_9 T_{3N}),$$
$$a_{13} = [\gamma_1 (T_{1M} - T_{1N})]/(2^N - 1),$$
$$a_{14} = [\gamma_2 (T_{2M} - T_{2N})]/(2^N - 1),$$
$$a_{15} = [\gamma_3 (T_{3M} - T_{3N})]/(2^N - 1),$$
$$a_{16} = \gamma_1 T_{1N} + \gamma_2 T_{2N} + \gamma_3 T_{3N}.$$

Здесь β_k и γ_k — постоянные преобразования координат цвета T_1, T_2, T_3 в равноконтрастные координаты цвета. Разложение функций u'', v'' и Y'' в ряд Тейлора в окрестности их среднего значения с сохранением в разложении членов до второго порядка включительно дает приближенно средние значения координат цветности и яркости восстановленного элемента изображения. Таким образом [1],

$$E\{u'' \mid M\} = \frac{a_1 E\{T'_{1T} \mid T_{1T}\} + a_2 E\{T'_{2T} \mid T_{2T}\} + a_3 E\{T'_{3T} \mid T_{3T}\} + a_4}{a_5 E\{T'_{1T} \mid T_{1T}\} + a_6 E\{T'_{2T} \mid T_{2T}\} + a_7 E\{T'_{3T} \mid T_{3T}\} + a_8} +$$
$$+ \frac{1}{2}\sigma^2\{T'_{1T} \mid T_{1T}\}\left[\frac{\partial^2 u''}{\partial T'^2_{1T}}\right] + \frac{1}{2}\sigma^2\{T'_{2T} \mid T_{2T}\}\left[\frac{\partial^2 u''}{\partial T'^2_{2T}}\right] +$$
$$+ \frac{1}{2}\sigma^2\{T'_{3T} \mid T_{3T}\}\left[\frac{\partial^2 u''}{\partial T'^2_{3T}}\right] \qquad (24.3.11)$$

[1]) Соотношение для v'' аналогично соотношению для u''.

и

$$E\{Y''\,|M\} = a_{13}E\{T'_{1T}\,|\,T_{1T}\} + a_{14}E\{T'_{2T}\,|\,T_{2T}\} + a_{15}E\{T'_{3T}\,|\,T_{3T}\} + a_{16}, \qquad (24.3.12)$$

где $E\{T'_{iT}|\,T_{iT}\}$ и $\sigma^2\{T'_{iT}\,|\,T_{iT}\}$ — условное среднее и условная дисперсия i-й компоненты на выходе канала связи — определяются выражениями (24.1.7) и (24.1.12).

Аналогично получаются выражения для условных дисперсий u'' и Y'':

$$\sigma^2\{u''\,|\,M\} \approx \sigma^2\{T'_{iT}\,|\,T_{iT}\}\left[\frac{\partial u''}{\partial T'_{1T}}\right]^2 + \sigma^2\{T'_{2T}\,|\,T_{2T}\}\left[\frac{\partial u''}{\partial T'_{2T}}\right]^2 + \sigma^2\{T'_{3T}\,|\,T_{3T}\}\left[\frac{\partial u''}{\partial T'_{3T}}\right]^2 \qquad (24.3.13)$$

и

$$\sigma^2\{Y''\,|\,M\} \approx a_{13}^2\sigma^2\{T'_{1T}\,|\,T_{1T}\} + a_{14}^2\sigma^2\{T'_{2T}\,|\,T_{2T}\} + a_{15}^2\sigma^2\{T'_{3T}\,|\,T_{3T}\}. \qquad (24.3.14)$$

24.3.2. ПЕРЕДАЧА ЯРКОСТИ И КООРДИНАТ ЦВЕТНОСТИ

Точные выражения для условных моментов распределений u'', v'' и Y'' при передаче яркости и координат цветности получаются таким же способом, какой использовался в предыдущем разделе. При прямом расчете по формуле

$$E\{(u'')^n\,|\,M\} = \sum_{t'_{1T}=0}^{2N-1}\ \sum_{t'_{2T}=0}^{2N-1}\ \sum_{Y'_T=0}^{2N-1} [u''\{t'_{1T},\ t'_{2T},\ Y'_T\}]^n \times P\{t'_{1T}\,|\,t_{1T}\}\,P\{t'_{2T}\,|\,t_{2T}\}\,P\{Y'_T\,|\,Y_T\} \qquad (24.3.15)$$

и по соответствующим формулам для $E\{(v'')^n\,|\,M\}$ и $E\{(Y'')^n\,|\,M\}$ величины u'', v'' и Y'' должны определяться для всех комбинаций принятых значений яркости и координат цветности. Так же, как и в предыдущем разделе, в случае канала с малой вероятностью ошибок можно перейти к приближенным выражениям. Операцией ограничения можно при этом пренебречь, выражая u'', v'' и Y'' непосредственно через принятые сигналы t_{iT}, Y. Таким образом,

$$u'' \approx (a_1t'_{1T} + a_2t'_{2T} + a_3)/(a_4t'_{1T} + a_5t'_{2T} + a_6), \qquad (24.3.16)$$

$$v'' \approx (a_7t'_{1T} + a_8t'_{2T} + a_9)/(a_4t'_{1T} + a_5t'_{2T} + a_6), \qquad (24.3.17)$$

$$Y'' \approx a_{10}Y'_T + a_{11}, \qquad (24.3.18)$$

где

$$a_1 = \alpha_1 (t_{1M} - t_{1N}),$$
$$a_2 = \alpha_2 (t_{2M} - t_{2N}),$$
$$a_3 = (\alpha_1 t_{1N} + \alpha_2 t_{2N} + \alpha_3)(2^N - 1),$$
$$a_4 = \alpha_4 (t_{1M} - t_{1N}),$$
$$a_5 = \alpha_5 (t_{2M} - t_{2N}),$$
$$a_6 = (\alpha_4 t_{1N} + \alpha_5 t_{2N} + \alpha_6)(2^N - 1),$$
$$a_7 = \alpha_7 (t_{1M} - t_{1N}),$$
$$a_8 = \alpha_8 (t_{2M} - t_{2N}),$$
$$a_9 = (\alpha_7 t_{1N} + \alpha_8 t_{2N} + \alpha_9)(2^N - 1),$$
$$a_{10} = (Y_M - Y_N)/(2^N - 1),$$
$$a_{11} = Y_N,$$

причем α_k — постоянные преобразования передаваемых сигналов в равноконтрастные координаты цвета. Следует отметить, что координаты цветности u'', v'' не зависят от принятого значения яркости Y'_T; что же касается значения Y'', то оно не зависит от принятых координат t'_{1T}, t'_{2T} только в отсутствие операции ограничения. Разложение в ряды Тейлора до членов второго порядка приводит к условным средним

$$E\{u'' \mid M\} = \frac{a_1 E\{t'_{1T} \mid t'_{1T}\} + a_2 E\{t'_{2T} \mid t_{2T}\} + a_3}{a_4 E\{t'_{1T} \mid t_{1T}\} + a_5 E\{t'_{2T} \mid t_{2T}\} + a_6} + \\ + \frac{1}{2}\sigma^2\{t'_{1T} \mid t_{1T}\}\left[\frac{\partial^2 u''}{\partial t'^2_{1T}}\right] + \frac{1}{2}\sigma^2\{t'_{2T} \mid t_{2T}\}\left[\frac{\partial^2 u''}{\partial t'^2_{2T}}\right] \quad (24.3.19)$$

и

$$E\{Y'' \mid M\} = a_{10} E\{Y'_T \mid Y_T\} + a_{11}, \quad (24.3.20)$$

а также к условным дисперсиям

$$\sigma^2\{u'' \mid M\} = \sigma^2\{t'_{1T} \mid t_{1T}\}\left[\frac{\partial u''}{\partial t'_{1T}}\right]^2 + \sigma^2\{t'_{2T} \mid t_{2T}\}\left[\frac{\partial u''}{\partial t'_{2T}}\right]^2 \quad (24.3.21)$$

и

$$\sigma^2\{Y'' \mid M\} = a_{10}^2 \sigma^2\{Y'_T \mid Y_T\}. \quad (24.3.22)$$

24.3.3. ЧИСЛЕННЫЕ ОЦЕНКИ

Расчет смещений яркости и цвета был произведен для следующих систем передаваемых координат: системы *RGB* основных цветов приемника НТСЦ, системы *YIQ* передаваемых сигналов НТСЦ, а также системы координат *UVW*, определяемой равноконтрастным цветовым графиком МКО (компонента *V* соответствует яркости).

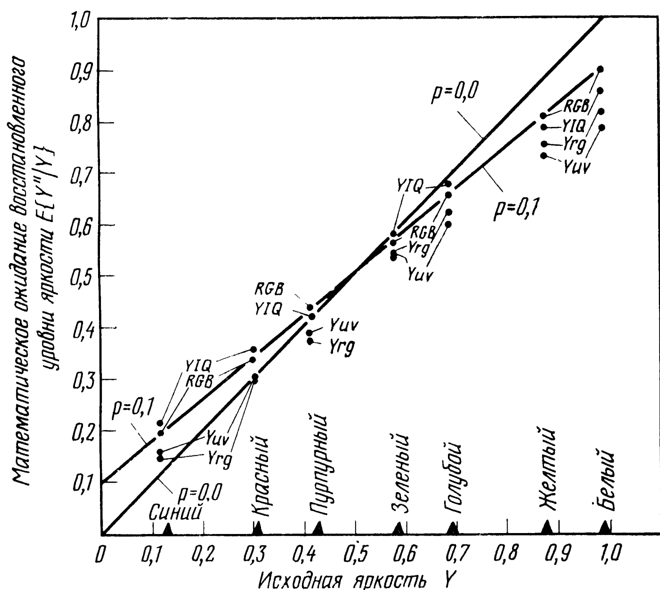
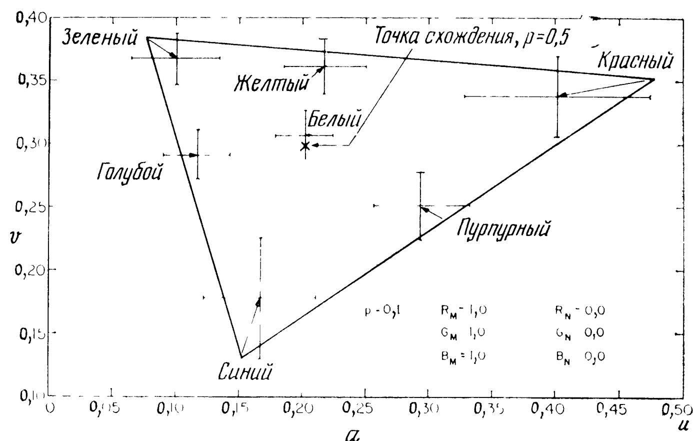


Рис. 24.3.2. Смещение восстановленного уровня яркости под воздействием помех при передаче цветного изображения по двоично-симметричному каналу.

На рис. 24.3.2 показана зависимость среднего смещения яркости для ряда испытательных цветов от фактического значения яркости при передаче в координатах (R, G, B), (Y, I, Q), (Y, r, g) и (Y, u, v). Предполагалось, что каждая компонента кодируется 6-разрядным двоичным числом, а вероятность ошибки при передаче по каналу связи $p = 0{,}1$. Данные, представленные в виде отдельных точек, получены непосредственным расчетом по формулам, аналогичным (24.3.7) и (24.3.15). Для сравнения на рис. 24.3.2 приведены также две сплошные линии, показывающие ожидаемое смещение яркости, приближенно определяемое формулой (24.2.5) в отсутствие операций ограничения при воспроизведении уровней красного, зеленого или синего. Приближенное равенство переходит в точное при передаче координат цвета R,

G, B, поскольку в этом случае сигналы не ограничиваются. В остальных случаях оказалось, что приближение сохраняет достаточную точность, если вероятность ошибки при передаче по каналу



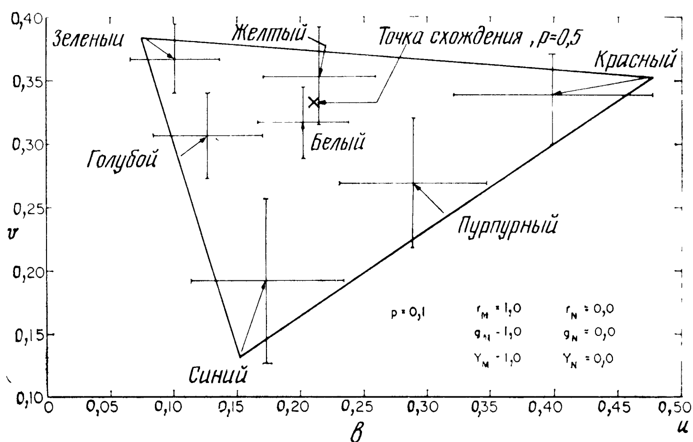
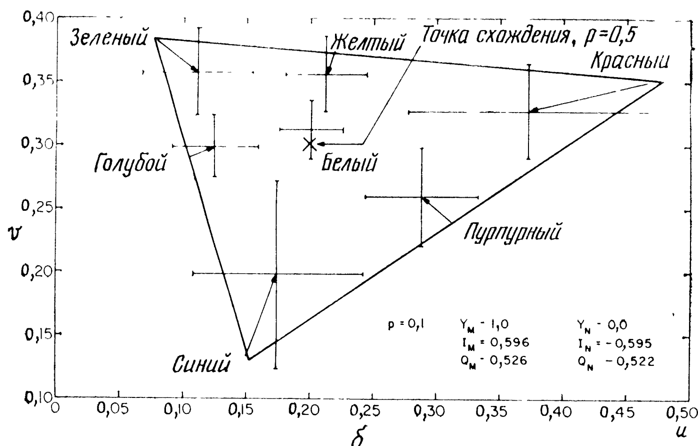


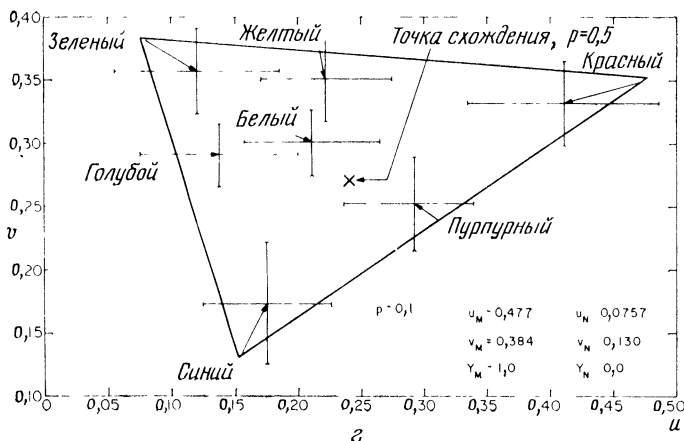
Рис. 24.3.3. Смещение координат цветности под воздействием помех при пере ческие ошибки воспроизведения цветов.)

а — передача сигналов R, G, B; *б* — передача сигналов Y, I, Q; *в* — передача сигналов

не превышает 0,001. Вообще, действие помех проявляется в смещении уровня яркости по направлению к середине диапазона ее воспроизводимых значений. Для большинства испытательных

цветов среднее смещение яркости при передаче координат цвета оказывается меньшим, чем при передаче координаты яркости и цветности.





даче по двоично-симметричному каналу. (Кресты определяют среднеквадрати-

Y, *r*, *g*; *г* — передача сигналов *Y*, *u*, *v*

Смещение координат цветности *u*, *v* под воздействием помех в канале связи иллюстрируется с помощью графиков на рис. 24.3.3. Данные относятся к семи испытательным цветам

и охватывают четыре рассмотренные выше системы передачи. Каждая из стрелок начинается в точке, указывающей передаваемый цвет, а кончается в точке, соответствующей восстанавливаемому в среднем цвету. Перекрещивающиеся отрезки прямых определяют ожидаемое среднеквадратическое отклонение цвета. Во всех этих примерах вероятность возникновения ошибки в канале составляет $p = 0{,}1$, а каждая из передаваемых компонент кодируется 6-разрядным двоичным числом. При передаче координат цвета R, G, B действие помех проявляется в снижении насыщенности переданного цвета и приближении его к белому цвету. Этот результат отвечает интуитивной оценке ситуации. В самом деле, все координаты цвета под воздействием помех будут в среднем одинаково приближаться к центру диапазона их изменения, т. е. к значению 0,5. В пределе с приближением вероятности ошибок p к значению 0,5 ожидаемая величина каждой из принятых координат R, G, B составит 0,5, какова бы ни была их переданная величина. Таким образом, восстановленный цвет станет серым при среднем уровне яркости. При передаче координат Y, I, Q помехи также проявляются в снижении насыщенности цвета в результате его приближения к белому цвету. По мере приближения p к значению 0,5 ожидаемые уровни принятых координат I и Q стремятся к нулю, так что восстанавливаемый цвет становится в пределе ахроматическим. Среднее смещение цвета при передаче сигналов Y, r, g направлено к желтоватому оттенку. Для объяснения этого эффекта рассмотрим опять предельную ситуацию при $p = 0{,}5$. Ожидаемый в пределе уровень для каждой из компонент r и g составляет 0,5. Следовательно, среднее значение координаты $b = 1 - r - g$, определяющей долю синего, стремится к нулю. Равные же количества красного и зеленого создают желтый цвет. Наконец, при передаче сигналов Y, u, v среднее смещение цвета направлено к пурпурному цвету. Это объясняется сильной зависимостью координат цветности u и v соответственно от красной и синей координат цвета. С повышением вероятности ошибок каждое из значений u, v приближается к 0,5, что влечет за собой увеличение доли красного и синего в восстановленном цвете и уменьшение доли зеленого.

Для проверки расчетных данных был проведен ряд экспериментов с помощью ЭВМ. Цветное изображение в цифровой форме подвергалось действию помех при передаче по двоично-симметричному каналу. Вероятность ошибок устанавливалась при этом на уровне $p = 0{,}01$. Среднее смещение цвета во всех экспериментах оказалось в согласии с расчетными данными, представленными на цветовых графиках. По субъективным оценкам, помехи менее всего заметны при передаче первичных координат цвета R, G, B. В этом случае среднее смещение цвета направлено в сторону серого, лишенного насыщенности цвета. В системах, использующих

для передачи другие координаты, среднее смещение цвета отклоняется от направления к серому в сторону главного дополнительного цвета.

24.4. ДЕЙСТВИЕ ПОМЕХ ПРИ ПЕРЕДАЧЕ ИЗОБРАЖЕНИЙ ПОСРЕДСТВОМ ДИКМ

Для восстановления уровня элемента в системе кодирования с предсказанием результат предсказания на приемной стороне, оценка уровня, складывается с принятым по каналу связи квантованным разностным сигналом — ошибкой предсказания. Возникновение ошибки при передаче по каналу приводит к искажению уровня при восстановлении очередного элемента, так что результаты предсказания на приемной стороне не будут совпадать с оценками уровня, вырабатываемыми в ходе предсказания на передающей стороне. Как следствие влияние всякой ошибки распрост-



Рис. 24.4.1. Воздействие помех канала связи на изображение в трехразрядной системе ДИКМ.

раняется на дальнейшие элементы по ходу строчной развертки до тех пор, пока не будет произведена коррекция или не произойдет случайной компенсации этого влияния последующей ошибкой в передаче. Чтобы ограничить распространение влияния ошибок в процессе кодирования с предсказанием, обычно производят периодический сброс системы, передавая непосредственно уровень начального элемента каждой строки с помощью ИКМ. Этот уровень используется как при передаче, так и на приемной стороне, аналогично результату предсказания.

Действие помех при передаче изображений системой ДИКМ с трехразрядной шкалой квантования разностного сигнала иллюстрирует рис. 24.4.1. Процесс передачи сопровождается сбросом системы в начале каждой строки для коррекции ошибок. При моделировании этого процесса на ЭВМ весовой коэффициент A_k, с которым учитывался в ходе предсказания предшествующий элемент строки, был принят равным 0,9. В результате такого



Рис. 24.4.2. Воздействие помех канала связи на изображение в случае трехразрядной системы кодирования с двумерным предсказанием 4-го порядка.

«интегрирования с утечкой» заметность полос ошибок несколько убывает по мере продвижения вдоль строки. Тенденция к ослаблению влияния ошибки по мере удаления от места ее возникновения сохраняется и при дальнейшем уменьшении весовых коэффициентов, однако это одновременно снижает способность системы к высококачественной передаче изображения в отсутствие помех.

Действие помех в случае кодирования с пространственным предсказанием 4-го порядка показано на рис. 24.4.2. Влияние ошибки распространяется как по вертикали, так и по горизонтали от места ее возникновения. Однако при одинаковой вероятности ошибок субъективное качество изображения значительно улучшается по сравнению с системой ДИКМ.

Для исправления ошибок в системах ДИКМ часто прибегают к периодической передаче предсказанных значений элементов, которые кодируются с помощью 8-разрядных комбинаций на передающей стороне; эти оценки используются на приемной стороне вместо соответствующих результатов предсказания [7]. Влияние

ошибки распространяется в этом случае не далее чем на период между проверками процесса предсказания в приемнике. При этом, однако, не следует упускать из виду возможность появления ошибок в проверочных 8-разрядных комбинациях. Другой подход к исправлению ошибок в системах ДИКМ состоит в последовательном сравнении значений разностных сигналов по мере их приема. Значительное расхождение между двумя последовательными значениями рассматривается при этом как свидетельство появления ошибки, и очередной разностный сигнал заменяется предшествующим. При этом исключаются разкие искажения уровня, но одновременно отбрасываются разностные сигналы, соответствующие реальным контурам, что снижает резкость восстановленного изображения.

24.5. ДЕЙСТВИЕ ПОМЕХ ПРИ ПЕРЕДАЧЕ ИЗОБРАЖЕНИЙ ПОСРЕДСТВОМ КОДИРОВАНИЯ НА ОСНОВЕ ПРЕОБРАЗОВАНИЯ

В случае кодирования на основе преобразования воздействие помех на передаваемое изображение носит иной характер, чем в системах ИКМ и ДИКМ. Поскольку для восстановления изображения используется взвешенная линейная комбинация коэффициентов преобразования, влияние ошибок в передаче коэффициентов усредняется для всех элементов восстановленного изображения. С другой стороны, ошибка хотя бы в единственном коэффициенте затрагивает все элементы изображения.

На рис. 24.5.1 показано действие помех на изображения, переданные с помощью зонального и порогового кодирования коэффициентов наклонного преобразования. Как и следовало ожидать, зональное кодирование значительно лучше противостоит вредному действию помех, чем пороговое кодирование с его чувствительностью к ошибкам в указании позиций коэффициентов, прошедших пороговый отбор.

24.6. СРАВНЕНИЕ РАЗЛИЧНЫХ МЕТОДОВ КОДИРОВАНИЯ

В части 6 этой книги описано большое число кодирующих систем. Всестороннее сравнение различных систем затруднительно, поскольку слишком много факторов определяют в каждом случае эффективность кодирования и сложность его практического осуществления. Совершенствование средств обработки и запоминания изображений привело в настоящее время к тому, что сложность осуществления кодирования перестала быть определяющим ограничением при разработке систем кодирования изображений. Стало возможным построение систем, требующих запоминания целого ряда кадров и достаточно сложной вычислительной об-

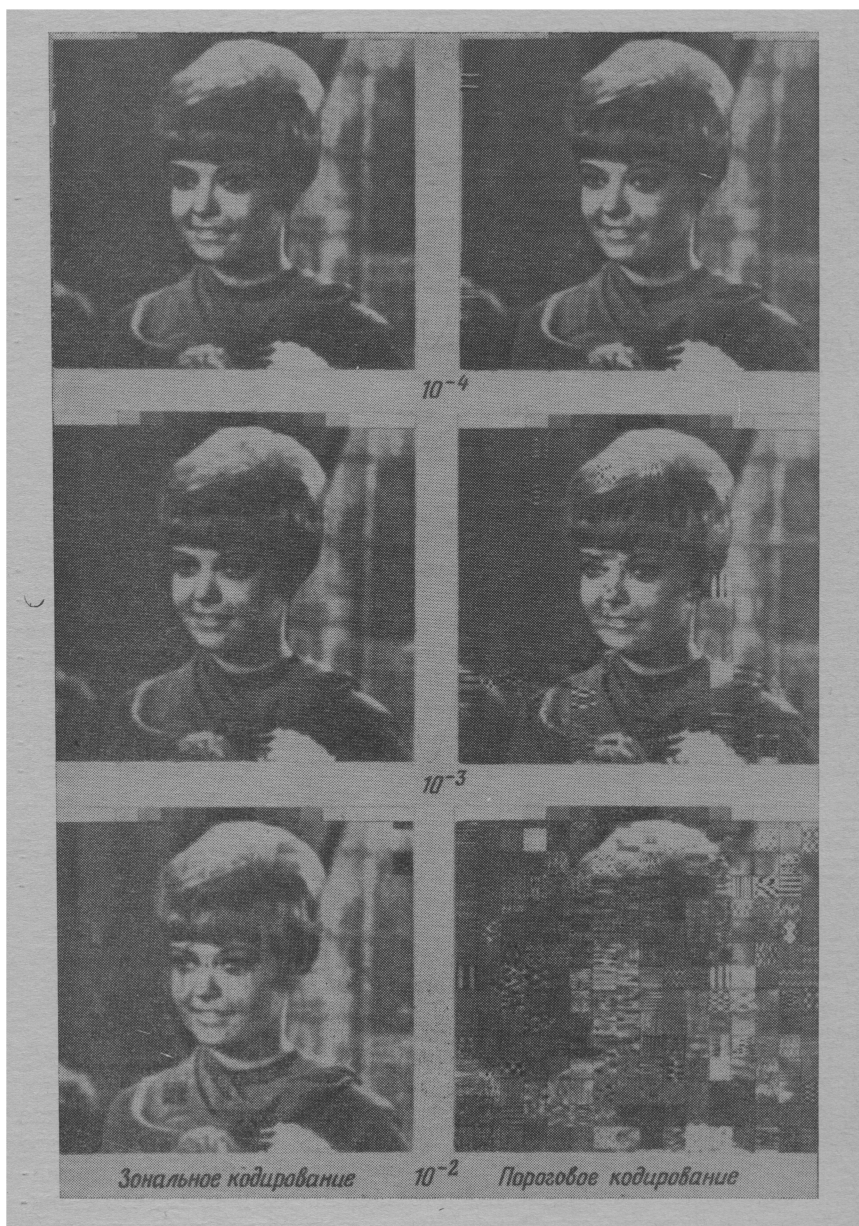


Рис. 24.5.1. Воздействие помех канала связи на изображение при кодировании с применением наклонного преобразования в случае затраты 1,5 дв. ед./эл. Размер блока 16×16 элементов.

работки данных, а также действующих в реальном масштабе времени телевизионной передачи. В связи с этим в данном разделе рассмотрение ограничивается оценкой эффективности кодирования в различных системах.

Диаграмма, представленная на рис. 24.6.1, позволяет сравнить ряд адаптивных и неадаптивных систем кодирования по такому показателю, как удельная затрата двоичных цифр. Как это обычно

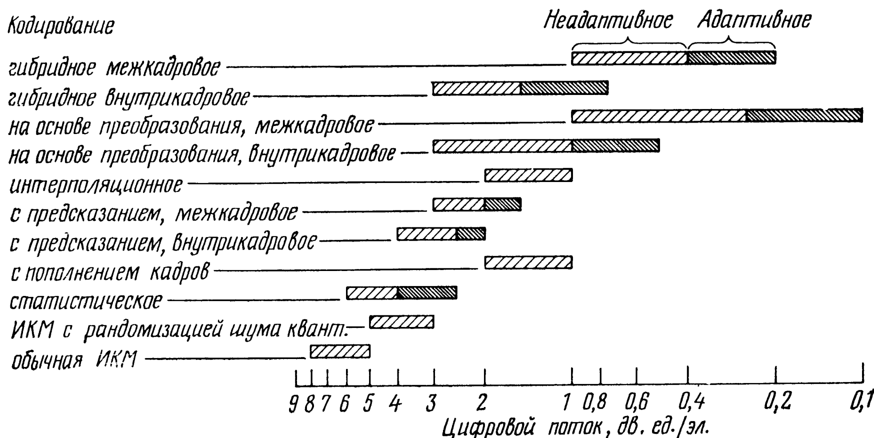


Рис. 24.6.1. Сравнение различных систем кодирования.

и предполагается, межкадровое кодирование имеет преимущество перед внутрикадровым, так же как адаптивное кодирование — перед неадаптивным. Это сравнение, конечно, не полно — диаграмма не содержит данных о погрешности кодирования. На рис. 24.6.2 показана теоретическая зависимость удельной затраты двоичных цифр от среднеквадратической ошибки кодирования для неадаптивных вариантов нескольких систем. Показана также граница, характеризующая предельные теоретические возможности двумерного кодирования. Для адаптивных процедур кодирования теоретические оценки, связывающие расход двоичных цифр с погрешностью, не получены. Диаграмма на рис. 24.6.2 построена в предположении, что источником сообщений служит марковский процесс с разделимой функцией корреляции, причем коэффициенты корреляции по горизонтальному и вертикальному направлениям, а также вдоль временной координаты, ρ_H, ρ_V и ρ_T, равны между собой и составляют каждый 0,95. Предполагается также, что каждая из кодирующих систем спроектирована с учетом статистических свойств источника и оптимальна в смысле минимума среднеквадратической ошибки кодирования. Основные характеристики этих систем приведены в табл. 24.6.1.

Таблица 24.6.1

Параметры систем кодирования

Система кодирования	Вероятностная модель	Размер блока	Метод кодирования
ИКМ	Распределение Рэлея	1×1	Равномерный код
Линейное предсказание	Распределение Лапласа	1×1	Равномерный код
Внутрикадровое кодирование с преобразованием	Распределения Рэлея (для постоянной составляющей) и Гаусса	16×16	Зональное кодирование, равномерный код
Гибридное внутрикадровое кодирование, ДИКМ-преобразование	Распределение Лапласа	1×16	Зональное кодирование, равномерный код
Межкадровое кодирование с преобразованием	Распределения Рэлея (для постоянной составляющей) и Гаусса	$16\times16\times16$	Зональное кодирование, равномерный код
Гибридное межкадровое кодирование, ДИКМ-преобразование	Распределение Лапласа	$16\times16\times1$	Зональное кодирование, равномерный код

В заключение следует отметить, что за несколько последних лет для передачи и хранения изображений разработан ряд методов кодирования, нашедших практическое применение. В результате теоретических исследований и субъективных испытаний накоплены данные, вполне достаточные для обоснованной сравни-

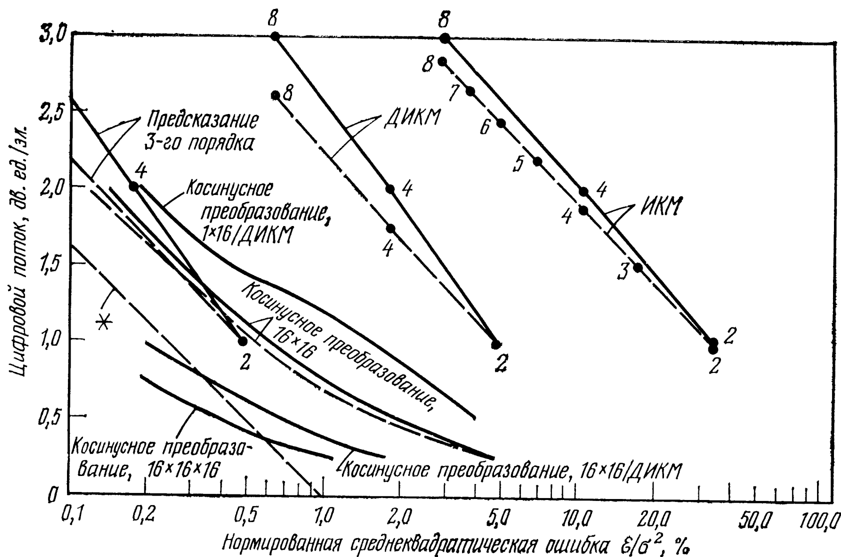


Рис. 24.6.2. Зависимость между эффективностью и среднеквадратической ошибкой кодирования для различных систем кодирования изображений.

— равномерный код; — — — статистический код; ● число уровней квантования; * граница эффективности двумерного кодирования, определяемая скоростью создания информации. $R = {}^1/_2 \log_2 \left[\frac{(1-\rho_H^2)(1-\rho_V^2)}{\mathscr{E}} \right]$.

тельной оценки конкурирующих систем кодирования с учетом сложности их практического осуществления и конкретных условий применения.

ЛИТЕРАТУРА

1. Huang T. S., Chikhaoui M. T., The Effect of BSC on PCM Picture Quality, *IEEE Trans. Inf. Theory*, **IT-13**, 2, 270—273 (April 1967).
2. Young I. T. Mott-Smith J. C., On Weighted PCM, *IEEE Trans. Inf. Theory* (*Correspondence*), **IT-11**, 4, 596—597 (October 1965).
3. Huang T. S., Tretiak O. J., Prasada B., Yamaguchi Y., Design Considerations in PCM Transmission of Low-Resolution Monochrome Still Pictures, *Proc. IEEE*, **55**, 3, 331—335 (March 1967). [Имеется перевод: Хуан, Третьяк, Прасада, Ямагуши. Соображения по конструированию КИМ-систем для передачи неподвижных одноцветных изображений низкой четкости. — ТИИЭР

тематический вып. «Сокращение избыточности», 1967, т. 55, № 3, с. 97—102.]

4. Pratt W. K., Digital Color Image Coding and Transmission, University of Southern California, Electronic Sciences Laboratory, USCEE Report 403, June 1971.
5. Knight J. M., Maximum Acceptable Bit Error Rates for PCM Analog and Digital TV Systems, Proceedings National Telemetering Conference, 1962, pp. 1—9.
6. Papoulis A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.
7. Arguello R. J., Sellner H. R., Stuller J. A., The Effect of Channel Errors in the Differential Pulse-Code Modulation Transmission of Sampled Imagery, *IEEE Trans. Commun. Tech.*, **COM-19**, 6, 926—933 (December 1971).
8. Katz R. L., Sciulli J. A., The Performance of an Adaptive Image Data Compression System in the Presence of Noise, *IEEE Trans. Inf. Theory*, **IT-14**, 2, 273—279 (March 1968).
9. Andrews H. C., Pratt W. K., Transform Image Coding, in: Computer Processing in Communications, Vol. XIX, Polytechnic Press, Brooklyn, New York, 1969, pp. 63—84.
10. Anderson G. B., Huang T. S., Errors in Frequency-Domain Processing of Images, Proceedings Spring Joint Computer Conference, Vol. 34, Boston, Massachusetts, 1969, pp. 173—185.

Приложение 1

ОСНОВНАЯ ЛИТЕРАТУРА

Книги

Ahmed N., Rao K. R., Orthogonal Transforms for Digital Signal Processing, Springer-Verlag, New York, 1975.

Andrews H. C., with contributions by Pratt W. K. and Caspari K., Computer Techniques in Image Processing, Academic Press, New York, 1970. [Имеется перевод: Эндрюс Г., Применение вычислительных машин для обработки изображений. — М.: Энергия, 1977.]

Andrews H. C., Introduction to Mathematical Techniques in Pattern Recognition, Wiley-Interscience, New York, 1972.

Andrews H. C., Hunt B. R., Digital Image Restoration, Prentice-Hall, Englewood Cliffs, N. J., 1977.

Born M., Wolf E., Principles of Optics, Pergamon Press, New York, 1970. [Имеется перевод: Борн М., Вольф Э., Основы оптики. — М.: Наука, 1973.]

Cathey W. T., Optical Information Processing and Holography, Wiley, New York, 1974.

Cheng G. C., Ed., Pictorial Pattern Recognition, Thompson, Washington, D. C., 1968.

Cornsweet T. N., Visual Perception, Academic Press, New York, 1970.

Duda R. O., Hart P. E., Pattern Classification and Scene Analysis, Wiley-Interscience, New York, 1973. [Имеется перевод: Дуда Р., Харт П., Распознавание образов и анализ сцен. — М.: Мир, 1976.]

Evans R. M., Hanson W. T., Brewer W. L., Principles of Color Photography, Wiley, New York, 1953.

Fink D. G., Ed., Television Engineering Handbook, McGraw-Hill, New York, 1957.

Fukunaga K., Introduction to Statistical Pattern Recognition, Academic Press, New York, 1972. [Имеется перевод: Фукунага К., Введение в статистическую теорию распознавания образов. — М.: Наука, 1979.]

Goodman J. W., Introduction to Fourier Optics, McGraw-Hill, New York, 1968. [Имеется перевод: Гудмен Дж., Введение в фурье-оптику. — М.: Мир, 1970.]

Grasselli A., Automatic Interpretation and Classification of Images, Academic Press, New York, 1969.

Graybill F. A., Introduction to Matrices with Applications in Statistics, Wadsworth, Belmont, Calif., 1969.

Harmuth H. F., Transmission of Information by Orthogonal Functions, 2nd ed., Springer-Verlag, New York, 1972. [Имеется перевод: Хармут X. Ф., Передача информации ортогональными функциями. — М.: Связь, 1975.]

Huang T. S., Tretiak O. J., Eds., Picture Bandwidth Compression, Gordon and Breach, New York, 1972.

Huang T. S., Eds., Topics in Applied Physics: Picture Processing and Digital Filtering, Vol. 6, Springer-Verlag, New Vork, 1975. [Имеется перевод: Хуанг Т. (ред.), Обработка изображений и цифровая фильтрация. — М.: Мир, 1979.]

Hunt R. W. G., The Reproduction of Colour, Wiley, New York, 1957.

Kaneff S., Ed., Picture Language Machines, Academic Press, New York, 1970.

Lipkin B. S., Rosenfeld A., Picture Processing and Psychopictorics, Academic Press, New York, 1970.

Mees C. E., The Theory of Photographic Process, Macmillan, New York, 1966.
Meisel W. S., Computer-Oriented Approaches to Pattern Recognition, Academic Press, New York, 1972.
Nevatia R., Structured Descriptions of Complex Curved Objects for Recognition and Visual Memory, Springer-Verlag, New York, 1977.
Oppenheim A. V., Schaefer R. W., Digital Signal Processing, Prentice-Hall, Englewood Cliffs, N. J., 1975. [Имеется перевод: Оппенгейм А. В., Шафер Р. В., Цифровая обработка сигналов. — М.: Связь, 1979.]
Papoulis A., Systems and Transforms with Applications in Optics, McGraw-Hill, New York, 1968. [Имеется перевод: Папулис А., Теория систем и преобразований в оптике. — М.: Мир, 1971.]
Preston K., Jr., Coherent Optical Computers, McGraw-Hill, New York, 1972.
Rabiner L. R., Gold B., Theory and Application of Digital Signal Processing, Prentice-Hall, Englewood Cliffs, N. J., 1975. [Имеется перевод: Рабинер Л., Гоулд Б., Теория и применение цифровой обработки сигналов. — М.: Мир, 1978.]
Rosenfeld A., Picture Processing by Computer, Academic Press, New York, 1969. [Имеется перевод: Розенфельд А., Распознавание и обработка изображений с помощью вычислительных машин. — М.: Мир, 1972.]
Rosenfeld A., Kak A. C., Digital Image Processing, Academic Press, New York, 1976.
Tippett J. T., et al., Eds., Optical and Electro-Optical Information Processing, MIT Press, Cambridge, Mass., 1965.
Winston P. H., Ed., The Psychology of Computer Vision, McGraw-Hill, New York, 1975. [Имеется перевод: Уинстон П. (ред.), Психология машинного зрения. — М.: Мир, 1978.]
Winston P. H., Artificial Intelligence, Addison-Wesley, New York, 1977. [Имеется перевод: Уинстон П., Искусственный интеллект. — М.: Мир, 1980.]
Wyszecki G., Stiles W. S., Color Science, Wiley, New York, 1967.

Специальные выпуски

Special Issue on Redundancy Reduction, *Proc. IEEE*, **55**, 3, (March 1967). [Имеется перевод: Сокращение избыточности. Тематический выпуск. — ТИИЭР, 1967, т. 55, № 3.]
Special Issue on Digital Communications, *IEEE Commun. Tech.*, **COM-19**, 6, Part I (December 1971).
Special Issue on Digital Picture Processing, *Proc. IEEE*, **60**, 7 (July 1972). [Имеется перевод: Обработка изображений при помощи ЦВМ. Тематический выпуск. — ТИИЭР, 1972, т. 60, № 7.]
Special Issue on Two-Dimensional Signal Processing, *IEEE Trans. Computers*, **C-21**, 7 (July 1972).
Special Issue on Digital Image Processing, *IEEE Computer*, **7**, 5 (May 1974).
Special Issue on Digital Signal Processing, *Proc. IEEE*, **63**, 4 (April 1975). [Имеется перевод: Цифровая обработка сигналов. Тематический выпуск. — ТИИЭР, 1975, т. 63, № 4.]

Обзорные работы

Andrews H. C., Tescher A. G., Kruger R. P., Image Processing by Digital Computer, *IEEE Spectrum*, **9**, 7, 20—32 (July 1972).
Andrews H. C., Digital Image Restoration: A Survey, *IEEE Computer*, **7**, 5, 36—45 (May 1975).
Habibi A., Robinson G. S., A Survey of Digital Picture Coding, *IEEE Computer*, **7**, 5, 21—34 (May 1974).

Huang T. S., PCM Picture Transmission, *IEEE Spectrum*, **2**, 12, 57—60 (December 1965).
Huang T. S., Image Enhancement — A Review, *Opto-Electronics*, **1** (February 1969).
Huang T. S., Digital Holography, *Proc. IEEE*, **59**, 9, 1335—1346 (September 1971). [Имеется перевод: Хуанг, Цифровая голография. — ТИИЭР, 1971, т. 59, № 9.]
Huang T. S., Schreiber W. F., Tretiak O. J., Image Processing, *Proc. IEEE*, **59**, 11, 1586—1609 (November 1971). [Имеется перевод: Хуанг, Шрайбер, Третьяк, Обработка изображений. — ТИИЭР, 1971, т. 59, № 11.]
Hunt B. R., Digital Image Processing, *Proc. IEEE*, **63**, 4, 693—708 (April 1975). [Имеется перевод: Хант, Цифровая обработка изображений. — ТИИЭР, 1975, т. 63, № 4.]

Приложение 2

ПРЕОБРАЗОВАНИЕ КООРДИНАТ ЦВЕТА И ЦВЕТНОСТИ

В системе с тремя основными цветами данный цвет может быть определен либо тремя координатами цвета T_1, T_2, T_3, либо координатами цветности t_1, t_2 и яркости Y. Если дано описание цвета в какой-либо системе координат, то можно перейти к любой другой системе координат.

1. *Преобразование координат цвета*

Пусть T_1, T_2, T_3 — координаты цвета в исходной системе координат. Тогда координаты цвета T_1, T_2, T_3 в новой системе определяются следующим образом:

$$\tilde{T}_1 = m_{11}T_1 + m_{12}T_2 + m_{13}T_3, \tag{1}$$

$$\tilde{T}_2 = m_{21}T_1 + m_{22}T_2 + m_{23}T_3, \tag{2}$$

$$\tilde{T}_3 = m_{31}T_1 + m_{32}T_2 + m_{33}T_3, \tag{3}$$

где m_{ij} — элементы матрицы преобразования. В табл. 1 приведены матрицы преобразования для основных систем координат цвета.

2. *Преобразования координат цвета в координаты цветности и яркости*

Пусть

$$t_1 \equiv T_1/(T_1 + T_2 + T_3), \tag{4}$$

$$t_2 \equiv T_2/(T_1 + T_2 + T_3) \tag{5}$$

и

$$\tilde{t}_1 \equiv \tilde{T}_1/(\tilde{T}_1 + \tilde{T}_2 + \tilde{T}_3), \tag{6}$$

$$\tilde{t}_2 \equiv \tilde{T}_2/(\tilde{T}_1 + \tilde{T}_2 + \tilde{T}_3) \tag{7}$$

Таблица 1

Матрицы преобразования координат цвета

Выходные координаты цвета	Входные координаты цвета R_c	G_c	B_c	R_N	G_N	B_N	X	Y	Z	U	V	W	Y	I	Q
R_c	1,000	0,000	0,000	1,167	−0,146	−0,151	2,365	−0,897	−0,468	2,846	0,507	−0,936	0,848	1,325	0,540
G_c	0,000	1,000	0,000	0,114	0,753	0,159	−0,515	1,426	0,089	−0,639	1,159	0,178	1,016	−0,265	−0,154
B_c	0,000	0,000	1,000	−0,001	0,059	1,128	0,005	−0,014	−1,009	1,521	−3,041	2,018	1,172	−1,247	1,688
R_N	0,842	0,156	0,091	1,000	0,000	0,000	1,910	−0,533	−0,288	2,432	0,332	−0,576	1,000	0,956	0,621
G_N	−0,129	1,320	−0,203	0,000	1,000	0,000	−0,985	2,000	−0,028	−1,519	2,083	−0,057	1,000	−0,272	−0,647
B_N	0,008	−0,069	0,897	0,000	0,000	1,000	0,058	−0,118	0,896	1,440	−2,823	1,803	1,000	−1,106	1,703
X	0,490	0,310	0,200	0,607	0,174	0,201	1,000	0,000	0,000	1,500	0,000	0,000	0,967	0,318	0,594
Y	0,177	0,813	0,011	0,299	0,587	0,114	0,000	1,000	0,000	0,000	1,000	0,000	1,000	0,000	0,000
Z	0,000	0,010	0,990	0,000	0,066	1,117	0,000	0,000	1,000	1,500	−3,000	2,000	1,173	−1,238	1,870
U	0,327	0,207	0,133	0,405	0,116	0,133	0,667	0,000	0,000	1,000	0,000	0,000	0,653	0,208	0,403
V	0,177	0,814	0,010	0,299	0,587	0,114	0,000	1,000	0,000	0,000	1,000	0,000	1,000	0,000	0,000
W	0,020	1,071	0,408	0,145	0,827	0,627	−0,500	1,500	0.500	0,000	0,000	1,000	1,599	−0,780	0,623
Y	0,178	0,818	0,016	0,299	0,587	0,114	0,000	1,000	0,000	0,000	1,000	0,000	1,000	0,000	0,000
I	0,540	−0,263	−0,174	0,596	−0,274	−0,322	1,407	−0,842	−0,451	1,403	0,534	−0,907	0,000	1,000	0,000
Q	0,246	−0,675	0,404	0,211	−0,523	0,312	0,932	1,189	0,233	1,757	−1,898	0,470	0,000	0,000	1,000

— координаты цветности в исходной и новой системах координат. Тогда из соотношений (1)—(3) следует, что

$$\tilde{t}_1 = \frac{\beta_1 T_1 + \beta_2 T_2 + \beta_3 T_3}{\beta_4 T_1 + \beta_5 T_2 + \beta_6 T_3}, \qquad (8)$$

$$\tilde{t}_2 = \frac{\beta_7 T_1 + \beta_8 T_2 + \beta_9 T_3}{\beta_4 T_1 + \beta_5 T_2 + \beta_6 T_3}, \qquad (9)$$

где

$$\begin{aligned} \beta_1 &\equiv m_{11}, \\ \beta_2 &\equiv m_{12}, \\ \beta_3 &\equiv m_{13}, \\ \beta_4 &\equiv m_{11} + m_{21} + m_{31}, \\ \beta_5 &\equiv m_{12} + m_{22} + m_{32}, \\ \beta_6 &\equiv m_{13} + m_{23} + m_{33}, \\ \beta_7 &\equiv m_{21}, \\ \beta_8 &\equiv m_{22}, \\ \beta_9 &= m_{23} \end{aligned} \qquad (10)$$

и m_{ij} — элементы матрицы преобразования координат цвета T_1, T_2, T_3 в координаты цвета $\tilde{T}_1$, $\tilde{T}_2$, $\tilde{T}_3$. Яркость Y связана с исходными координатами цвета следующим соотношением:

$$Y = n_{21} T_1 + n_{22} T_2 + n_{23} T_3, \qquad (11)$$

где n_{ij} — элементы матрицы преобразования исходных координат T_1, T_2, T_3 в координаты X, Y, Z.

3. *Преобразование координат цветности и яркости в новые координаты цветности и яркости*

Подставив формулы

$$T_1 = t_1 (T_1 + T_2 + T_3), \qquad (12)$$

$$T_2 = t_2 (T_1 + T_2 + T_3), \qquad (13)$$

$$T_3 = (1 - t_1 - t_2)(T_1 + T_2 + T_3) \qquad (14)$$

в выражения (8) и (9), получим

$$\tilde{t}_1 = \frac{\alpha_1 t_1 + \alpha_2 t_2 + \alpha_3}{\alpha_4 t_1 + \alpha_5 t_2 + \alpha_6}, \qquad (15)$$

$$\tilde{t}_2 = \frac{\alpha_7 t_1 + \alpha_8 t_2 + \alpha_9}{\alpha_4 t_1 + \alpha_5 t_2 + \alpha_6}, \qquad (16)$$

где

$$\begin{aligned}
\alpha_1 &\equiv m_{11} - m_{13},\\
\alpha_2 &\equiv m_{12} - m_{13},\\
\alpha_3 &\equiv m_{13},\\
\alpha_4 &\equiv m_{11} + m_{21} + m_{31} - m_{13} - m_{23} - m_{33},\\
\alpha_5 &\equiv m_{12} + m_{22} + m_{32} - m_{13} - m_{23} - m_{33},\\
\alpha_6 &\equiv m_{13} + m_{23} + m_{33},\\
\alpha_7 &\equiv m_{21} - m_{23},\\
\alpha_8 &\equiv m_{22} - m_{23},\\
\alpha_9 &\equiv m_{23}
\end{aligned} \qquad (17)$$

и m_{ij} — элементы матрицы преобразования координат T_1, T_2, T_3 в координаты $\tilde{T}_1$, $\tilde{T}_2$, $\tilde{T}_3$.

4. Преобразование координат цветности и яркости в координаты цвета

В общем случае, когда исходные координаты цветности не являются координатами цветности МКО, преобразование выполняется в два этапа. Из формул (1)—(3) следует, что

$$\tilde{T}_1 = n_{11}X + n_{12}Y + n_{13}Z, \qquad (18)$$

$$\tilde{T}_2 = n_{21}X + n_{22}Y + n_{23}Z, \qquad (19)$$

$$\tilde{T}_3 = n_{31}X + n_{32}Y + n_{33}Z, \qquad (20)$$

где n_{ij} — элементы матрицы преобразования системы координат XYZ в заданную систему координат $\tilde{T}_1\tilde{T}_2\tilde{T}_3$.

Найдем теперь координаты X, Y, Z. Из формул (15) и (16) следует, что

$$x = \frac{\alpha_1 t_1 + \alpha_2 t_2 + \alpha_3}{\alpha_4 t_1 + \alpha_5 t_2 + \alpha_6}, \qquad (21)$$

$$y = \frac{\alpha_7 t_1 + \alpha_8 t_2 + \alpha_9}{\alpha_4 t_1 + \alpha_5 t_3 + \alpha_6}, \qquad (22)$$

где постоянные α_k можно получить из коэффициентов m_{ij} формулы (17) для преобразования системы координат $T_1T_2T_3$ в систему координат XYZ. Так как

$$X = (x/y)\,Y, \tag{23}$$

$$Z = [(1 - x - y)/y]\,Y, \tag{24}$$

то координаты X и Z, входящие в формулы (18)—(20), определяются следующим образом:

$$X = \frac{\alpha_1 t_1 + \alpha_2 t_2 + \alpha_3}{\alpha_5 t_1 + \alpha_6 t_2 + \alpha_7}\,Y, \tag{25}$$

$$Z = \frac{(\alpha_4 - \alpha_1 - \alpha_5)\,t_1 + (\alpha_5 - \alpha_2 - \alpha_6)\,t_2 + (\alpha_6 - \alpha_3 - \alpha_7)}{\alpha_5 t_1 + \alpha_6 t_2 + \alpha_7}\,Y. \tag{26}$$

Приложение 3

СТАТИСТИЧЕСКОЕ КОДИРОВАНИЕ

В этом приложении описываются два основных метода статистического кодирования: Шеннона—Фано и Хаффмена.

Рис. 1 иллюстрирует метод кодирования Шеннона—Фано. Согласно этому методу, символы, образующие исходное сообщение, вместе с их вероятностями появления представляются в виде таблицы в порядке убывания вероятностей. Затем символы делятся на две группы с наиболее близкими друг к другу суммарными вероятностями. Первым разрядам кодовых слов первой группы присваивается значение нуль, а второй группы — единица. В результате последующего деления групп символов на подгруппы определяются следующие разряды кодовых слов. Средняя длина кодовых слов и энтропия источника сообщений указаны на рисунке.

Символы	Вероятность	Кодовое слово
a_1	0,30	00
a_2	0,25	01
a_3	0,25	10
a_4	0,10	110
a_5	0,10	111

$L_c = (2)\,[0{,}3 + 0{,}25 + 0{,}25] + (3)\,[0{,}1 + 0{,}1] = 2{,}2$ дв. ед.,

$H = -(0{,}3)\log_2(0{,}3) - (2)\log_2(0{,}25) - (2)(0{,}1)\log_2(0{,}1)$,

$H = 2{,}185$ дв. ед.,

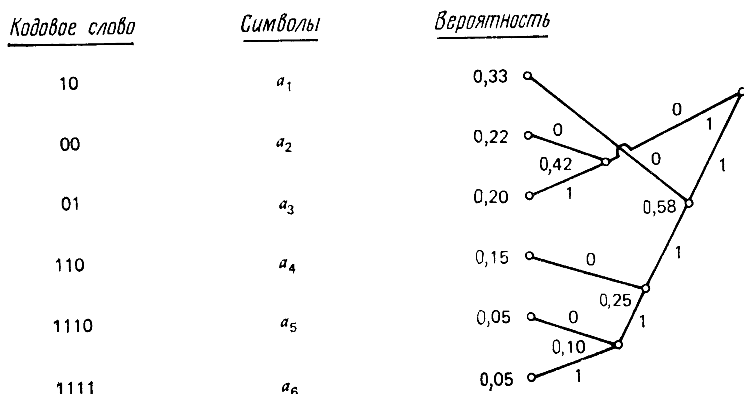
$E_F = L_c/H = 0{,}993$.

Рис. 1. Пример кодирования методом Шеннона—Фано.

Метод кодирования Хаффмена поясняется примером, приведенным на рис. 2. Два символа с наименьшими вероятностями объединяются в узел кодового дерева. Вероятности символов суммируются и приписываются узлу. Затем объединяются следующие символы или узлы с наименьшей вероятностью, как показано на рисунке. Этот процесс продолжается, пока ветви дерева не сойдутся к одному узлу — вершине. Ветви дерева, сходящиеся к узлу, обозначаются нулем и единицей. Чтобы получить кодовое

слово, соответствующее данному символу, идут по дереву в обратном направлении от вершины к символу, записывая нули или единицы, которыми обозначены пройденные ветви.

Для получения кодов Шеннона—Фано и Хаффмена вместо графических методов, требующих построения кодового дерева, можно воспользоваться алгебраическими методами. Обычно



$$L_c = (2)\,(0{,}33 + 0{,}22 + 0{,}20) + (3)\,(0{,}15) + (4)\,(0{,}05 + 0{,}05),$$

$L_c = 2{,}35$ дв. ед.,

$$H = -(0{,}33)\log_2(0{,}33) - (0{,}22)\log_2(0{,}22) - (0{,}20)\log_2(0{,}20) -$$
$$- (0{,}15)\log_2(0{,}15) - (2)(0{,}05)\log_2(0{,}05),$$

$H = 2{,}32$ дв. ед.,

$E_F = L_c/H = 0{,}897.$

Рис. 2. Пример кодирования методом Хаффмена.

алгоритмы кодирования Шеннона—Фано и Хаффмена приводят к одинаковой средней длине кодового слова. Для некоторых наборов вероятностей сообщений средняя длина кодового слова достигает значения энтропии источника сообщений. Можно показать, что код Хаффмена всегда дает минимальную среднюю длину кодового слова.

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

Книги

Cappellini V., Constantinides A., Eds., Digital Signal Processing, Academic Press, London, 1980.
Huang T. S., Ed., Topics in Applied Physics: Two-Dimensional Digital Signal Processing I. Linear Filters, vol. 42, Springer-Verlag, New York, 1981.
Huang T. S., Ed., Topics in Applied Physics: Two-Dimensional Digital Signal Processing II. Transforms and Median Filters, vol. 43, Springer-Verlag, New York, 1981.
Judd D., Wyszecki G., Color in Biseness, Science and Industry, Wiley, New York, 1975. [Имеется перевод: Джад Д., Вышецки Г. Цвет в науке и технике: Пер. с англ./Под ред. Л. Ф. Артюшина. — М.: Мир, 1978.]
Kunt M., de Coulon F., Eds., Signal Processing: Theories and Applications, North-Holland Publ. Comp., Amsterdam, 1980.
Stucki P., Eds., Advances in Digital Image Processing, Plenum Press, New York, 1979.

Специальные выпуски

Special Issue on Digital Image Processing, *IEEE Commun. Tech.*, **COM-26**, 5 (October 1978).
Special Issue on Digital Image Processing, *Proc. IEEE*, **69**, 5 (May 1981). [Имеется перевод: Цифровая обработка изображений. Тематический выпуск. — ТИИЭР, 1981, т. 69, № 5.]

Обзорные работы

Netrevali A., Limb J. O., Picture coding: A Review., *Proc. IEEE*, **68**, 3, 366—406 (March 1980). [Имеется перевод: Нетревали, Лимб. Кодирование изображений. Обзор. — ТИИЭР, 1980, т. 68, № 3.]
Jain A. K., Image Data Compression: A Review, *Proc. IEEE*, **69**, 3, 349—389 (March 1981). [Имеется перевод: Джайн, Сжатие видеоинформации. Обзор. — ТИИЭР, 1981, т. 69, № 3.]

РАБОТЫ СОВЕТСКИХ СПЕЦИАЛИСТОВ

Книги

Антипин М. В. Интегральная оценка качества телевизионных изображений. — Л.: Наука, 1970.
Василенко Г. И. Голографическое опознавание образов. — М.: Сов. радио, 1977.
Василенко Г. И. Теория восстановления сигналов. — М.: Сов. радио, 1979.
Завалишин Н. В., Мучник И. Б. Модели зрительного восприятия и алгоритмы анализа изображений. — М.: Наука, 1974,
Игнатьев Н. К. Дискретизация и ее приложения. — М.: Связь, 1980.
Ковалевский В. А. Методы оптимальных решений в распознавании изображений. — М.: Наука, 1976.
Кустарев А. К. Колориметрия цветного телевидения. — М.: Связь, 1967.
Лебедев Д. С., Цуккерман И. И. Телевидение и теория информации. — М.: Госэнергоиздат, 1965.

Лебедев Д. С. (ред.). Иконика. — М.: Наука, 1968.
Лебедев Д. С. (ред.). Иконика. Пространственная фильтрация изображений. Фотографические системы. — М.: Наука, 1970.
Лебедев Д. С. (ред.). Иконика. Обработка изображений. Цифровая голография. — М.: Наука, 1975.
Цуккерман И. И., Кац Б. М., Лебедев Д. С и др. Цифровое кодирование телевизионных изображений. — М.: Радио и Связь, 1981.
Ярославский Л. П., Мерзляков Н. С. Методы цифровой голографии. — М.: Наука, 1977.
Ярославский Л. П. Введение в цифровую обработку изображений. — М.: Сов. радио, 1979.

Специальные выпуски

Иконика. Цифровая обработка и фильтрация изображений. Проблемы кибернетики, вып. 38, М.: ВИНИТИ, 1978.
Иконика — новое направление в изучении изображений. Труды ГОИ, том 44, вып. 178, Л.: 1979.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ

ЧАСТЬ 4. РЕСТАВРАЦИЯ И УЛУЧШЕНИЕ ИЗОБРАЖЕНИЙ

ЧАСТЬ 5. АНАЛИЗ ИЗОБРАЖЕНИЙ

УВАЖАЕМЫЙ ЧИТАТЕЛЬ!

Ваши замечания о содержании книги, ее оформлении, качестве перевода и другие просим присылать по адресу: 129820, Москва, И-110, ГСП, 1-Рижский пер., 2, издательство «Мир».

Уильям Прэтт

ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

Старший научный редактор Л. П. Якименко
Младший научный редактор Е. П. Орлова
Художник В. И. Харламов
Художественный редактор Л. Е. Безрученков
Технический редактор Н. Д. Толстякова
Корректор В. С. Соколов

ИБ № 2371

Сдано в набор 25.03.81.
Подписано к печати 19.11.81.
Формат $60 \times 90^1/_{16}$.
Бумага типографская № 2.
Гарнитура литературная. Печать высокая.
Бум. л. 15,00 Усл. печ. л. 30,0. Усл. кр. отт. 30,0.
Уч-изд. л. 28,18. Изд. № 20/1652.
Тираж 10 000 экз. Зак. 597. Цена 2 р. 20 к.

ИЗДАТЕЛЬСТВО «МИР»
129820, Москва, И-110, ГСП, 1-й Рижский пер., 2

Ленинградская типография № 6 ордена Трудового
Красного Знамени Ленинградского объединения
«Техническая книга» им. Евгении Соколовой
Союзполиграфпрома при Государственном комитете СССР
по делам издательств, полиграфии и книжной торговли
193144, г. Ленинград, ул. Моисеенко, 10

2р. 20к.



У. Прэтт

Цифровая обработка изображений

2